AGENT-BASED SIMULATION OF ECOLOGICAL ALCOHOL SYSTEMS

by

Yasmin Hassan Said
A Dissertation
Submitted to the
Graduate Faculty
of
George Mason University
in Partial Fulfillment of
The Requirements for the Degree
of
Doctor of Philosophy
Computational Sciences and Informatics

Committee:

_________________________ Director: Edward J. Wegman

_________________________ David S. Anderson

_________________________ Jeffrey L. Solka

_________________________ William F. Wieczorek

_________________________ David Wong

_________________________ Jim X. Chen

_________________________ Dept. Chairperson: Edward J. Wegman

_________________________ Program Director: Peter A. Becker

_________________________ Menas Kafatos, Dean
Computational Sciences & Informatics

Date: _________________________ Spring Semester 2005
George Mason University
Fairfax, VA

Agent-Based Simulation of Ecological Alcohol Systems

A dissertation submitted in partial fulfillment of the requirements for the degree of Doctorate of Philosophy at George Mason University

by

Yasmin Hassan Said
Doctorate of Philosophy
George Mason University, 2005

Director: Bernard J. Dunn Professor Edward J. Wegman,
Data Sciences Program, School of Computational Sciences

Spring Semester 2005
George Mason University
Fairfax, VA

Dedication

This dissertation is dedicated to Dr. Edward J. Wegman, my academic advisor who gave me the strength and support that was needed in every stage of this process. I would also like to dedicate this dissertation to my parents and family for their encouragement and personal sustenance they so lovingly provided me.

Acknowledgements

Above all, I would like to acknowledge the assistance of my academic advisor, Dr. Edward J. Wegman for his continuous support and confidence in me during the hardest of times. I also acknowledge my advisory committee for their critical feedback that only served to strengthen my dissertation. Finally, I would like to thank my parents and family for all of their support and patience throughout the years and in dealing with my nonstop rambling of alcohol and alcoholics.

Table of Contents

List of Tables

List of Figures

Abstract

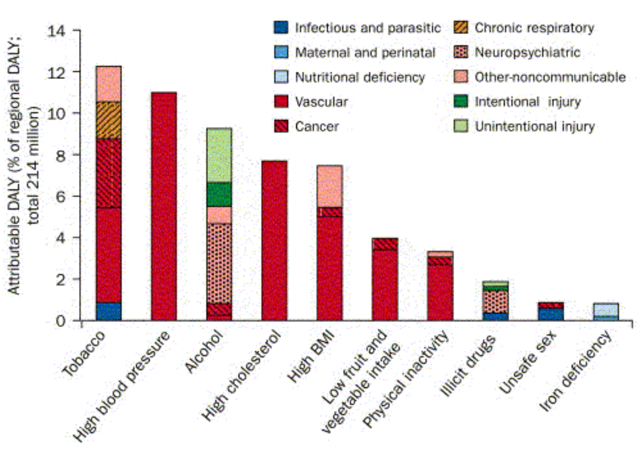
AGENT-BASED SIMULATION OF ECOLOGICAL ALCOHOL SYSTEMS

Yasmin Said, Ph.D.

George Mason University, 2005

Dissertation Director: Dr. Edward J. Wegman, The Bernard J. Dunn Professor of Information Technology and Applied Statistics, Data Sciences Program, School of Computational Sciences, George Mason University

This is research is to establish a modeling framework for alcohol abuse that allows evaluation of interventions meant to reduce adverse effects of alcohol overuse without the financial, social and other costs of imposing interventions that are ultimately ineffective (or even simply not cost effective). The framework is ecological (individual agents and their activities are represented), stochastic (neither individual behavior nor consequences of interventions are certain) and flexible. Constructing the framework involves interactions among the



Source: Ezzati, Majid, et al. "Selected major risk factors and global and regional burden of disease." *The Lancet.* 360.9343 (2002): 1347-1360.

domain science of alcohol studies, statistics, and computer science.

In the developed world such as the United States, based on Disability Adjusted Life Years (DALY) lost (Ezzati et al., 2002), alcohol is the third leading cause of mortality and morbidity. Much of the mortality and morbidity is associated intentional acts (for example violence, suicide) and unintentional injuries (for example, crashes, falls, burns) as a result of drinking behavior. Interventions designed to reduce both chronic and acute mortality and morbidity associated with alcohol use may be effective in limited circumstances. However, often suppressing one negative outcome can have the impact of increasing other negative outcomes. For example, adding police patrols to drinking hotspots may suppress assaults, but may increase driving under the influence and domestic violence. This research initiates a systemic approach to understanding the complex interactions and consequently exploring the effect of interventions within sectors of the complex ecological-like system that surrounds users and abusers of alcohol. The tool that I have developed effectively allows the investigator to play the “what if” games to find improved strategies and subsequently design intervention strategies toward accomplishing the desired reductions in negative outcomes. Such a strategy is particularly relevant to both the entire society and subsegments of it, e.g. the military because of the potential for alcohol abuse among young, unattached males and the resulting loss of military readiness, order and warfare capabilities (Dunham, 2004). The leap from genomic, microbiologic and neuronal understanding of the effects of alcohol to the behavior of individuals as influenced by alcohol is one that has not been thoroughly understood in the medical and psychological literature (Leonard and Blane, 1999). I believe that an ecological approach leads to a better understanding of the alcohol behavior on a macro-scale and allows insight into the linkage between the neuronal scale and the behavioral scale with a direct impact on treatment strategies.

There can be a large benefit to society in general from a reduction in the overall prevalence of acute outcomes related to alcohol use and abuse. Indeed, illicit drugs, a well-recognized problem, are only eighth in loss of DALY. Indeed, the US alcohol DALY loss is ranked third and is five times large that the illicit drug DALY loss. Ironically, the research funding associated with drug abuse research is five times that associated with alcohol research. I believe that the proposed research could have a strong societal impact. I have developed an agent-based stochastic directed graph (digraph) model using Fairfax County, VA.

Fairfax contains Fort Belvoir and the preliminary work already shows a higher than usual pattern of violence associated with alcohol use in this sector of Fairfax County. Thus, I have developed the concept within an interesting environment that contains civilian, military, and university student populations. While not fully developed as subpopulations in this model, this county serves as a test bed.

Chapter 1: Ecological Alcohol Systems

Ethanol or ethyl alcohol or simply alcohol is an addictive drug (Gilmore, 2004) sold in the United States that is both widely available legally without a prescription (Wikipedia, 2005). In relatively low dosages, it creates a euphoric sense of well being (Pharmacology Central, 2000) that often seduces the user into consuming higher dosages (Alcohol Problems in Later Life, 1998; Leonard and Blane, 1999). At higher dosage levels, alcohol is a depressant (University of Ottawa, 2005; Rutgers University, 2004; Australian Institute of Criminology, 2005) that suppresses both cognitive and motor functions in the brain (Fogarty, 2002; Fillmore, 1998). Because alcohol suppresses cognitive function, it impairs judgment (Nahas, 1997; Changeux, 2001). This can lead to a range of acute outcomes (Reiss, 1994) and violence (Moss, 1993; Huntingdonshire District Council, 2005) that includes assault and battery, suicide, sexual assault, murder, domestic violence (Butler Center for Research Hazelden Foundation, 2000; U.S. Department of Health and Human Services and SAMHSA's National Clearinghouse for Alcohol and Drug Information, 1995), and child abuse (NIAAA, 2005; Iowa Substance Abuse Information Center, 2005; Gmel and Rehm, 2003). Because alcohol suppresses motor function, driving while under the influence can lead to automobile crashes (Harrison, 1997) including fatal crashes, another acute outcome of alcohol usage. Of course at the highest dosage levels, alcohol poisoning can kill the alcohol user (Asylum Seeker Resource Centre, 2005). Lethal dose 50 (LD50[1]) for alcohol occurs at 0.35 Blood Alcohol Concentration (BAC) (Canadian Centre for Occupational Health and Safety, 1999, The American Heritage Dictionary of the English Language, 2004; The American Heritage Stedman's Medical Dictionary, 2002; answers.com, 2005). Clearly, interventions that mitigate the undesirable acute outcomes need to be explored, but they are often based on an incomplete understanding of the entire system of alcohol usage.

The study of alcohol usage and its effects can be addressed at different scales. The broadest understanding comes from studying the societal dynamics surrounding alcohol use. There is a complex alcohol system to study. The alcohol system is not unlike a classical ecological system[2]. The alcohol system involves the complex interactions in time and place among users (including casual drinkers, heavy users/alcohol abusers, binge drinkers, under age or young drinkers, and alcoholics), their family and peers, non-users, producers and distributors of alcohol products, law enforcement, courts, prevention activities, and treatment centers (NIAAA, 2000). The alcohol system must be understood

[1] Lethal dose 50 is that dosage of a substance that causes ½ the exposed population to die. This could be for a medication such as chemotherapy drugs, a toxic poison such as ricin, drugs such as alcohol or heroin, or an everyday substance such as drain cleaner. LD 30 would of course by analogy cause 30% of the exposed population to die.

[2] A classical ecological system is a collection of multiple organisms, their environment, their relationships, and their interactions.

also in terms of sub-populations and geo-spatial interactions among the diverse communities. In short, understanding the alcohol system involves many if not most of the same issues and the same level of complexity that face ecologists in understanding conventional ecological systems (Wilson and Dufour, 2000).

Because the consequences of alcohol misuse are so severe (including violence and life threatening situations) for individuals as well as for society, a tool that provides policy insights into the effectiveness of interventions is most desirable. It is the purpose of this dissertation research to explore the development of such a tool. I would like to be able to answer “what if” questions, e.g. Is it more useful to intervene geo-spatially, within work places, within schools, within ethnic groups, at certain times of the day, with what age groups, with additional police, with additional taxes, with more severe penalties, with fewer alcohol outlets, with subsidies to treatment facilities? What is the effect of increasing populations? What is the effect of a shift in ethnic balance and with increased diversity? What would happen if previously abstaining populations began drinking? A whole host of policy issues could potentially be addressed with such a tool.

The development of such a tool involves extensive data collection and a serious mathematical modeling effort involving scenario development including mobility, work and leisure practices, and an understanding of the fabric of the social network. Such an effort is likely to uncover required data that are difficult to obtain or not yet collected. This project would be a multiple year effort beyond the scope of a single dissertation. In my dissertation, I have developed a proof of concept. By this I mean that this dissertation is a pilot study based on data collected from public sources. The effort is intended to illustrate that my approach can model the gross features of the ecological alcohol system, but to the extent that all of the needed data were not available, this dissertation cannot be regarded as the public policy tool that I eventually hope to create.

The expectation is that a well-calibrated model of the ecological alcohol system would enable simulations yielding insight into interventions likely to be successful in mitigating the acute effects of alcohol usage. In the research reported in this dissertation, I limit the scope of the modeling and simulation to a specific geographic region and to a short-term. However, I have in mind expanding this concept to include longitudinal data spanning longer time frames. Longer-term modeling is needed to account for adaptations of users to interventions and for aging effects in the population. Longer-term modeling increases complexity because more actors in the social network must be accounted for. I believe that this approach provides the foundation for more elaborate models addressing longer time durations and broader geographic regions. However, it will take considerably more effort in both model development and data collection to calibrate the model. The specifics of the approach and the scope of effort are described below.

The development of more elaborate models will constitute a considerable challenge in terms of model specificity, data collection, and model calibration. Much effort went into the data collection for the relatively simple model illustrated here. The better the data the fewer assumptions need to be made in specifying model probabilities. Complex models are not likely to capture all the feedback loops. When interventions are made, unanticipated consequences may arise and new feedback loops corresponding to the unanticipated consequences may emerge from obscurity. This dissertation is focused on a proof of concept and the model presented here is not intended as the final model.

1.1 Ethanol, Ethyl, Grain Alcohol, Alcohol

Ethanol is a clear liquid with a fairly sweet taste in dilute solutions, but can result in a burning taste at higher concentrations. Ethanol, CH_3CH_2OH, is classified as an alcohol, which is characterized by a hydroxyl group attached to a carbon atom. Although the term alcohol originates from the Arabic al-kuhul, or a fine powder used as eye makeup, medieval alchemists later applied the word to products of distillations, which is where the current term gets its usage (Petrucci, 2001; Shakhashiri. 2005).

Physical properties of ethanol include a melting point of –114.1°C, a boiling point of 78.5°C, and a density of 0.789 grams/milliliter at 20°C. Relative to mercury with a freezing point of –40°C, alcohol is the fluid of choice in thermometers with low temperature readings and for use as antifreeze in automobile radiators (Petrucci, 2001; Shakhashiri, 2005).

From ancient times through today, ethanol has been produced through the fermentation process of sugars. The enzyme that provides the force in the conversion of simple sugars to ethanol and carbon dioxide is zymase, which is derived from yeast, through the following reaction:

$$C_6H_{12}O_6 \longrightarrow 2\ CH_3CH_2OH + 2\ CO_2$$

Although the production of ethanol through this fermentation reaction with the use of impure yeast cultures results in impurities such as glycerine and other organic acids, it is precisely these impurities that provide the flavors for beverage alcohol. Starches supplied from corn, potatoes, and other wheat plants including barley can also yield ethanol through fermentation, although starches must be broken down into simple sugars before the fermentation process can begin through the naturally occurring enzyme diastase. This enzyme is released through the germination of barley, which is therefore required to be the first step in producing alcohol from starchy plants (Petrucci, 2001; Shakhashiri, 2005).

Zymase is only active as an enzyme in ethanol concentrations up to 14 percent, at which point the enzyme is destroyed and fermentation is ceased. Pure ethanol cannot be produced by distillation because although ethanol is normally concentrated by the distillation of aqueous solutions, the constitution of vapor from aqueous ethanol results in four percent of the solution being water. To produce pure ethanol, dehydrating agents can be used to absorb any excess water, however, commercial ethanol is sold as 95 percent by volume ethanol, with the remaining five percent as water.

To prevent its consumption, industrial ethanol is often denatured and small amounts of poisonous or unpleasant substances are added to it. The price industrial corporations would have to pay in order to remove these substances would exceed the federal excise tax on alcoholic beverages, however, because some industries require undenatured ethanol, federal supervision in such cases is mandatory.

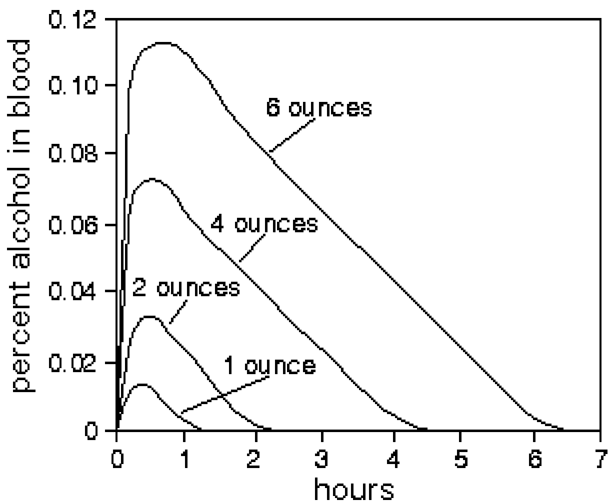
After swallowing an alcoholic beverage, ethanol is rapidly absorbed in the small intestine and distributed throughout the body entering body tissues in direct proportion with their water content. This results in more ethanol being distributed to the blood and brains rather than muscles or fat tissue. Because ethanol is significantly diluted by body fluids, a one-ounce shot of 100 proof distilled spirits, which is composed of a half of an ounce of ethanol, is diluted by a factor of 5000 in a 150-pound human, resulting in an approximate blood alcohol concentration of 0.02 percent.

As ethanol is a toxic substance, upon its consumption the body disposes of it immediately through alcohol dehydrogenase in the liver. Alcohol dehydrogenase converts ethanol into acetylaldehyde. As acetaldehyde is also a toxic substance, aldehyde dehydrogenase immediately converts acetylaldehyde into acetate ions.

$$CH_3CH_2OH \longrightarrow CH_3-\overset{\overset{\displaystyle O}{\|}}{C}-H + 2\ H$$

$$CH_3-\overset{\overset{\displaystyle O}{\|}}{C}-H + H_2O \longrightarrow CH_3-\overset{\overset{\displaystyle O}{\|}}{C}-O^- + 3\ H$$

Hydrogen atoms in the above equations become bound to nicotinamide-adenine dinucleotide (NAD) and results in the generation of NADH. In order for ethanol elimination to continue, NADH must constantly be converted back to NAD. As shown in the figure, blood alcohol levels change over time depending on the amount of ethanol consumed, with lower amounts of ethanol being cleared from the body more quickly (Petrucci, 2001; Shakhashiri, 2005).



Source: http://scifun.chem.wisc.edu/CHEMWEEK/ETHANOL/ethanol.html

Figure 1: A response curve showing how the body clears different amounts of alcohol in the blood.

The central nervous system is significantly affected by the consumption of ethanol. Instead of affecting muscles of the body, ethanol's effects are directed at the brain. Ethanol can serve as a depressant or general anesthetic because of its suppression of certain brain functions. However, ethanol consumed at low doses can suppress some inhibitory brain functions and can therefore act as a stimulant. With higher concentrations of ethanol, more brain functions are suppressed and reaction time becomes slowed down along with slurred speech, and other symptoms of intoxication. At especially high concentrations within the body, ethanol serves as a general anesthetic.

Because directly measuring alcohol levels in the brain is rather difficult, monitoring of blood alcohol level serves as a good substitute. Impairment of brain functions for most people begin to become noticed at around a blood alcohol percentage of 0.05, while clearly noticeable physical impediments become visible at around a percentage of 0.10, and slurred speech becomes evident at 0.15 percent. Most people lose consciousness at a blood alcohol level of 0.4 percent, and at a level of 0.5 percent the brain’s breathing center and the pumping of the heart can become anesthetized resulting in their impediment. While reaching such an extreme degree of intoxication is quite improbable, a 150 pound human may attain this level after the quick ingestion of a fifth of a gallon of 100 proof alcohol.

1.2 An Overview of the Overall Program

I developed a simulation model of the alcohol system based on the concept of a stochastic directed graph (Snijders, 2001). The details are described in the model approach chapter. The research discussed in this dissertation is intended as a first step in a larger research agenda I intend to pursue. Thus, as mentioned earlier, the present research is intended as a proof of concept. The concept is that relatively homogeneous clusters of people are identified along with their daily activities. In the parlance of social networks, *agents* simulate people. The activities are characterized by different states in the directed graph, and decisions resulting in actions by an agent move the agent from state to state in the directed graph. The leaf nodes in the graph represent a variety of outcomes, some of which are benign, but a number of which are acute alcohol-related outcomes (Dawson, 2000). Specifically, what is in mind is studying simultaneously the following acute outcomes: 1) assault and battery, 2) suicide, 3) domestic violence, 4) child abuse, 5) sexual assault, 6) murder, 7) DWI (with motor vehicle crashes including crashes that result in fatalities). The agents have probabilities associated with their transit from state to state through the directed graph. The structure of the directed graph and the associated probabilities are assessed based on national and local data and expert opinion. While not presently implemented in this dissertation, in order to increase the richness of the possible behaviors for the agents from a given cluster, a hierarchical Bayesian structure (De Freitas, et al., 2000) could be introduced and is planned for later work.

As agents are introduced into the directed graph model, their outcomes whether benign or acute accumulate so that a (multinomial) probability distribution can be estimated. The ultimate goal is to create a tool that will be useful for public policy formulation by allowing the analyst to investigate potential effects of interventions. I conceive this investigation as a two-part strategy. At a technical level, my tool allows the adjustment of one or more conditional probabilities that effectively alter the structure of the directed graph with the goal of assessing how those adjustments affect the probability distribution over the outcomes. It is possible that an intervention may reduce the incidence of one acute outcome, but increase the incidence of other acute outcomes. For example, reducing assaults at an off-license selling alcohol by increasing police patrols may increase the occurrence of DWIs and domestic violence because the user must leave the area to consume the purchases made at the off-license. The second part of the strategy is to develop interventions that will achieve the probability adjustments leading to a favorable reduction in the probabilities associated with acute outcomes. For example, focusing on reductions of drinking behaviors within certain populations or certain geographic regions could reduce the overall probability of acute outcomes. The goal is to study the alcohol system as a whole in order to evaluate best interventions for reducing the overall incidence of acute outcomes. As a policy tool, this will be helpful because the current non-systematic approach yields very limited capabilities in this regard.

The target experimental site is chosen as Fairfax County in Northern Virginia. There are several reasons for such a selection. There are readily identifiable subpopulations within Fairfax County including subpopulations that exhibit problem drinking behaviors. These include university and high school age populations, military populations, white-collar and blue-collar workers, and significant immigrant communities. In addition, there is significant local expertise in alcohol studies and access to data on alcohol use in this geographic region from public records and from alcohol related surveys. Local experts also have access to and experience with the Virginia Department of Alcoholic Beverage Control (http://www.abc.state.va.us, 2005), the Virginia Alcohol Safety Action Program (http://vasap.state.va.us, 2005), and other remediation and treatment programs listed in more detail in the subsequent chapter.

I believe that the three major thrusts of this research are

i) Model Building,

ii) Data Collection to Calibrate the Model,

iii) Computational Aspects of the Simulation Process.

The longer-term research should have sufficiently rich detail to assess the impact of interventions, but my present research is limited in scope consistent with a proof of concept. In particular, I limit the model to a single-day simulation, which defers considerations of producer's advertising, judicial outcomes, and treatment programs to a longer-term, larger-scale project. The focus on a single day is a result of many limiting factors, including the limited scope of a dissertation project, availability of data, the nature of what is doable with unfunded research, and regulations related to human subjects. I believe that the basic structure of my model is scalable to a future larger project.

1.3 Anticipated Major Innovations

There is a critical need to develop innovative approaches that can move the alcohol field forward in quantum leaps rather than small incremental steps. The need is highlighted in the World Health Organization's *The World Health Report 2002: Reducing Risks, Promoting Healthy Lives* (World Health Organization, 2002), which compares preventable risk factors for premature death and disability. In developed countries, such as the US, alcohol is the third leading cause of premature death and disease, exceeded by only tobacco and hypertension. Unfortunately, the current state of knowledge regarding interventions and their efficacy for reducing alcohol-related mortality and disability is relatively limited. Current approaches based on field trials and clinical research yield incremental information slowly. The development of the directed graph model offers the opportunity to move the alcohol field forward at a highly accelerated pace by

i) allowing simulations of the impact of known interventions

ii) stimulating thinking about innovative concepts for interventions, and

iii) using the results of the simulations to conduct targeted field studies of the promising interventions identified through the simulations.

The stochastic directed graph model involving agents from diverse communities represents a comprehensive view of a very important segment of the alcohol system and offers a dynamic simulation view of the system with simultaneous assessment of a variety of acute outcomes. This approach incorporates factors such as geographic distribution and operating hours of alcohol outlets, including bars, state liquor stores, grocery and convenience stores (selling beer and wine), and restaurants, also incorporate mobility information. Intervention procedures such as changes in alcohol service and distribution policies, law enforcement procedures, judicial sentencing policies, taxing policies and other government legislation can be assessed by altering the group specific probability structures. I believe that the most important innovation is the possibility of assessing simultaneously the effect of interventions on the probability distributions of acute outcomes with a dynamic simulation model.

Chapter 2: The Model

2.1 Stochastic Directed Graphs

I have developed a simulation model based on the concept of directed graphs. A directed graph (often called a digraph for brevity's sake) G is a pair (V,E) where V is a set of elements called vertices or nodes and E is a subset of the set of all ordered pairs (a,b), where a and b are vertices. (An element of E is called an edge or an arc of G). The pair (a,b) is not the same as the pair (b,a). Typically I regard the direction of the edge (a,b) as flowing from a to b. Conventionally an edge of the digraph is represented as two points representing a and b with an arrow whose tail is at a and whose head is at b. A graphic example with four vertices labeled a,b,c,d and four edges is given in Figure 2.

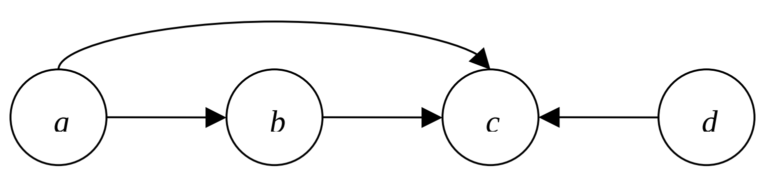


Figure 2: A simple digraph with four vertices and four directed edges.

More precisely, I model the alcohol system with an agent-dependent, time dependent stochastic digraph. The notion is that the vertices of the digraph represent the state of the agent (including such factors as physical location, present activity, and level of BAC) and the edges represent a decision/action that takes the agent into a new state. The agent represents any individual in the population including the alcohol users as well as the non-users. The edge going from one state to another has a conditional probability attached to it, hence the notion of a stochastic digraph. The conditional probability attached to a given edge depends on the specific sub-population, from which the agent is drawn, hence is agent-dependent, and, in the larger model, the conditional probability also depends on the time of day, hence it is time-dependent. In general, for this proof-of-concept modeling and simulation effort, I focus on short-term modeling of a single day. In essence, this removes the need to model court action, but not law enforcement; it removes the need to model the impact of producers, but not distributors; and it removes the need to model treatment centers, but not prevention programs. Clearly the limitation to one day simplifies the modeling process and the data requirements, but also weakens the achievable flexibility of such an approach. However, I am only trying to develop a proof of concept in this dissertation. All of these subtleties can be modeled ultimately. This approach allows flexibility in modeling so that more complex or less complex models may be formulated depending on the purpose and goal of a study. Modeling the more complex interactions introduces a need for putting feedback into the system and, of course, collecting additional data for estimating the conditional probabilities built into the model. Limiting the scope is consistent with demonstrating the proof of concept and does not reduce the generality of the proposed approach.

In general, the directed graph is tree-structured with leaf nodes (vertices with edges pointed to them, but for which there are no edges pointed away from them). The leaf nodes correspond to the outcomes, which may be one of the acute outcomes or may be outcomes that are completely benign. (In the current model, an agent can experience only one acute outcome per day. If the agent experiences no acute outcomes in the day, the agent has a benign outcome.) A path through the digraph represents the path of an agent from the agent's initial state to final resolution of his or her decisions/actions at the leaf node. The path may be very simple involving only two or three vertices or may be very complex, following the many activities of an agent throughout the entire day. Because there are conditional probabilities attached to each edge, the outcome for agents with identical starting states may be very different. In my planned follow-on research (described in Chapter 7) using longer durations, the role of courts and the penalties they impose, the role of treatment centers, whether treatment is successful or not, and the re-entry of offenders into old or new patterns of behavior can also be modeled. This will in general introduce feedback loops into the digraph, but will not alter the basic modeling-simulation technique.

I believe that the use of a directed graph is a very fruitful device in order to stimulate clear thinking about the sequence of states and actions possible for any agent. By creating the directed graph, one can sequence the set of states such that the consequences of a single decision/action can be carefully envisioned independent of the agent. The agent may affect the conditional probabilities (including possibly setting some to zero), but not the general structure of the digraph. In general, interventions alter the conditional probabilities differentially for agents from different sub-populations, socioeconomic status, geographic regions, age, and racial/ethnic backgrounds. Because the model and simulation account for a wide variety of acute and benign outcomes simultaneously, one is able to estimate probabilities of outcomes simultaneously, i.e. estimate probability distributions over the leaf nodes.

This is in contrast with the more conventional pure statistically based alcohol studies, which examine static data and draw limited conclusions. For example, most alcohol consumption is associated with consumption of beer and outlets that sell beer in large quantities tend to have the beer consumed nearby. Policing these outlets by additional visibility of law enforcement or increase in taxes, may reduce the level of violence locally, but there is no way to assess the impact on other acute outcomes or judge whether or not such actions reduce the overall probability of acute outcomes. The agent-based stochastic digraph model/simulation allows for the dynamic adjustment of conditional probabilities so that the final distribution of probabilities over all outcomes may be assessed. In general one would like to raise the probability of a benign outcome and simultaneously lower the probability of acute outcomes, such as assault, domestic violence or DWI. However, it is possible that certain interventions will reduce probabilities associated with some acute outcomes, but increase the probabilities associated with other acute outcomes, and, in fact, not reduce the overall probability of acute outcomes.

The stochastic digraph model can be exploited as a simulator by using Monte Carlo simulation to decide a path through the digraph for each agent generated. At each node (state) there is a conditional probability distribution for the next node (state) associated with the agent and time of day and the decision/action that the agent makes. The same decision/action for the same time of day and the same agent can have possibly different outcomes. However it must have some outcome so that the conditional probabilities must add to one. That is to say, at each node a randomly generated number between zero and one may be used to decide to which node the agent is taken. This ultimately leads each agent to transition to one leaf node, which may be an acute outcome or a benign outcome. By introducing many agents into the digraph, one can dynamically simulate the probability distribution of outcomes. By adjusting the interventions, which corresponds to adjusting certain transition probabilities, one can examine how those interventions affect the probability distributions of outcomes.

2.2 Activity Generation

My digraph model is closely related to models used to generate activities as part of more comprehensive transportation models. A prime example is the TRANSIMS system a comprehensive model developed under federal support at Los Alamos National Laboratory and currently commercialized by IBM. The TRANSIMS activity generator constructs activities, their locations, and method of travel between locations, for each member of every household simulated household in a city. The program is currently being used in a demonstration project in Portland, Oregon, a city whose greater metropolitan area has around 1.6 million residents.

Transportation planners regularly conduct extensive travel surveys, which are available for use. A prime example with which one is familiar is the Portland Travel/Activity Survey of 1994-1995. Participants in the survey kept diaries and recorded every activity with duration and location for a two-day period for each member of the survey households. In addition, extensive demographic information was recorded. There were 4,451 households and 10,048 individuals in the survey; 129,188 activities were recorded. It should be noted that drinking behavior was not recorded in this survey. A number of activities, presumably including drinking, were aggregated in a leisure activity category. However, the survey does provide valuable data on daily patterns of activity, periods of work, meals, and shopping and leisure activities. Based on these skeletal activity patterns, it is believed that it would be possible to use other data sources to build drinking behavior into these activity patterns.

The Metropolitan Council of Governments (COG) in the Washington, D.C. area, conducted a similar survey in 1992-93. Unfortunately, this survey is only a travel survey, not an activity survey. It only lists activities where travel was necessary. Other activities such as activities at home were not recorded. However, I eventually intend to use the COG survey in conjunction with the more detailed Portland survey. My expectation is to be able to use Portland and other comparable cities' activities as surrogates for activities of Fairfax County residents. However, this was not done as part of my dissertation research.

2.3 Why Northern Virginia?

The activity generator creates agents whose demographics and activities are representative of the population being modeled. I model the population of the county of Fairfax in Northern Virginia. This region is not a monolith and can serve as a model of the microcosm of diversity found throughout the nation. The choice of this population seems to be a good choice because of the relatively easy availability of appropriate data concerning alcohol behavior and because of interesting identifiable subpopulations residing in the area. The population of Fairfax County is quite affluent in general, and includes geographically identifiable sub-populations. In particular, George Mason University is a large urban university with a diverse student population approaching 30,000. Typically alcohol usage in such a setting is associated with binge drinking. Fairfax also contains Fort Belvoir, a large U.S. Army installation along the Route 1 corridor, which represents both a military population as well as a relatively less affluent population. There is also a comparatively large Hispanic population, notably from El Salvador, principally employed as laborers and in service industries.

The choice of the Northern Virginia is also advantageous from the point of view that Virginia has a tightly controlled system for distribution of distilled spirits, the Department of Alcoholic Beverage Control, through which policy interventions may be made. In contrast, beer and wine are sold at a large variety of outlets. Virginia has relatively strict enforcement of DWI offenses. Conviction of a third DWI is a felony in Virginia and everyone convicted of a DWI is automatically remanded to the Alcohol Safety Action Program (ASAP). The implication of this is that aggregated data on DWI convictions as well as data on the distribution of distilled spirits are relatively easy to gather.

Because of the demographic identifiability, general homogeneity, and geographic localization of the stratified subpopulations, it is practical to create in a limited time-frame an accurate generator of agents and their activities that will be representative of the entire population of these counties. I stratify the population demographically and by geographic location. Thus agents are chosen proportionally to the size of the population stratum and given representative activities associated with that stratum.

2.4 Some Cautions and Limitations of the Approach

Given the limited scale of the research in this dissertation, many issues, some potentially important, are necessarily not addressed. While I have attempted to develop some understanding of how significantly the omissions affect the simulation, it is likely that this understanding is incomplete.

Conceptually, perhaps the most glaring omission is that interaction among agents is not modeled explicitly (as it is, for example, in the vehicle movement module of TRANSIMS). As noted earlier, consumption of alcohol is, in some instances, a highly social activity, and it is clear that some adverse outcomes (for example, fights and automobile accidents) entail interaction among agents. A model that represents interactions explicitly is simply out of the question as an entry point. Instead, use of Bayesian hierarchical models introduces, but in a controlled fashion, dependence among agents that can serve as a surrogate for interactions.

Second, my approach is better at representing “quantitative” rather than “qualitative” changes in interventions. This is because the interventions are represented as changes in transition probabilities. Thus, for example, an increased level of law enforcement is easier to accommodate than a completely new kind of enforcement (such as devices that disable cars whose drivers are determined to be under the influence of alcohol).

Third, I have worked with existing data, which, because of funding available and limited availability of other resources, are not always the data I really needed. This is necessitated by the exploratory nature of my research. In the long run, my modeling framework helps to define the data needs.

Although complex, the stochastic digraph model implicitly has a Markov character to it. That is, once in a state, the transition to the next state does not depend on the entire history, but only on the current state. This is a simplifying assumption. It is mitigated, however, by the modeling of characteristics of agents, which implicitly models additional behavioral characteristics. I believe that this simplifying assumption is necessary in a complex model and will not substantially alter conclusions.

The Portland travel survey revealed that there is tremendous variability in work schedules, and in fact most people do not work eight-to-five shifts. There are roughly as many part-time workers as full-time, there are some workers beginning their shift at virtually any time of day, and workers go from one job to another. The implication is that a carelessly developed pure Markov model for scheduling work based solely on observed marginal probabilities would generate an appreciable proportion of agents working impossible (such as 24-hour) schedules.

These limitations notwithstanding, I believe that my approach is sufficiently rich to demonstrate proof of concept. The limitations would then be addressed in subsequent research. Indeed, to attempt to remove the limitations now would negate any possibility of proof of concept within the scale of this dissertation.

Chapter 3: The Data

3.1 At the Macro Level

A broad array of data sources are necessary to facilitate the development of the alcohol ecosystem model. These data are required for a number of purposes ranging from providing detailed population characteristics for the areas to be modeled to information on specific drinking behaviors by age and demographic groups. The availability and access to requisite data enhanced the feasibility of the project.

Demographic Information: US census data provide detailed information on the demographic distributions of characteristics such as age, gender, race/ethnicity, and socioeconomic status (e.g., median income, poverty status). These data are available from the decennial census and its updates. Data on all full-count and long form census items are typically available at the tract and block group levels of geography.

Alcohol-related Behaviors: These data are critical to the development of the digraph model because specific inputs are needed for the model and data on alcohol-related outcomes are used for model calibration. Local and state databases provide some of the relevant information; but no single data source could provide the detailed data necessary. Although no single source of data is sufficient; clearly, there are adequate data from a combination of local, county, state, national, and specialty data sources.

These data sources were supplemented by national databases including those presented and developed by NIAAA's Alcohol Epidemiologic Data System (AEDS). AEDS issues special reports on such topics as alcohol problem indicators and trend in alcohol-related mortality. There are a multitude of national data sources that were also utilized including the CDC’s Behavioral Risk Factor Surveillance System, National Survey on Drinking and Driving Attitudes and Behaviors, National Longitudinal Alcohol Epidemiologic Survey (NLAES), National Alcohol Surveys (conducted by the Alcohol Research Group), National Health Interview Survey (NHIS), National Health and Nutrition Surveys (NHANES), National Survey on Drug Use and Health (NSDUH, formerly NHSDA), and the National Survey of Substance Abuse Treatment Services (N-SSATS). This combination of data sources provides a varied and rich source of information for model building.

3.2 At the Micro Level

The focus on Fairfax in Northern Virginia was a deliberate choice not only because of advantages mentioned above. The required demographic and geographic data are available from county sources. Northern Virginia (metropolitan Washington, DC) transportation is extensively studied and an excellent database exists. Alcohol usage data are available from both the Virginia Department of Alcoholic Beverage Control and the Virginia Alcohol Safety Action Program. Information on acute outcomes involving felonies is part of the record of the Circuit Court (19th Judicial District).

3.3 Data Collection Description

I expended an enormous amount of effort in the data collection process. I contacted many government and private agencies to obtain the appropriate data needed for a thorough understanding and analysis.

I have been in contact with and the current data that have been collected were from the Virginia Department of Motor Vehicles (DMV), Virginia Department of Alcoholic Beverage Control, Virginia Police Department, Fairfax County Police Department, Fairfax County Crime Data Analysis Department, Fairfax County Criminal Investigation Bureau, Virginia Department of Health, Hospitals, INOVA Fairfax Hospital, INOVA Trauma Center, Fairfax/Falls Church Community Services Board (CSB), Office of Substance Abuse Services (OSAS), SAMHSA, Virginia Commonwealth Tax Administration Office, Fairfax County Board of Supervisors, U.S. Postal Address Management Services, Fairfax County Health Information Services, Fairfax/Falls Church Community Services Board Alcohol Drug Services, Division of Alcohol and Drug Services, Virginia Health Statistics, Fairfax County Citizen Assistance and Information, Fairfax County Demographic Information, Fairfax County Electoral Board, Fairfax County Geographic Information Services (GIS), Fairfax County Public Affairs, Fairfax County Maps and Publications Office, Fairfax County Department of Management and Budget, Census Bureau, and Bureau of Labor Statistics.

3.3.1 Synopsis of Data Currently Assembled

I have been able to collect an array of data. The following is a brief synopsis of the data that I have collected.

- I was able to obtain all zip codes within Fairfax County. There were 47 zip codes within Fairfax County. Some zip codes crossover to other counties. I obtained the percent of zip code addresses in Fairfax County (see Table 4 for more details). The percent of zip code in a given region was obtained from both the U.S. Census Bureau and the Address Management Office. In this research, the geographic definition and boundary for Fairfax County is the same as the definition for the region that is used by the U.S. Census Bureau, Address Management Office, and the Fairfax County GIS. For example, Fairfax City, the City of Falls Church, Alexandria City are not included in the Fairfax County. Our data are shown in Table 4 of Appendix B.

- The population and demographic information for all 47 zip codes was obtained. This data are presented by zip code region in Table 5 in Appendix B.
- ABC distilled spirits sales in Fairfax County for 2000-2004, broken down by each store in Fairfax County was obtained. There are 24 ABC stores in Fairfax County. A subset of the stores sales data are shown in Table 6 in Appendix B.
- Alcohol Establishment license information was also obtained. This information contains establishment names and addresses, along with associated company names, license status, and contact information. A small sample subset from this data source (which contains many thousands of records) is shown in Table 7 in Appendix B
- The number of alcohol establishments in each zip code, and the number of ABC stores also in each zip code was obtained. Based on these figures, I determined if the alcohol availability [based on the number of these locations] within that zip code was **Low** (0 to 15), **Moderate** (16 to 49), or **High** (greater than or equal to 50). This information is used later by the Alcohol Tree simulation to determine “hot spot” regions of intense acute outcomes. Further information about how region intensity is calculated is available in Section 4.2.2 Part 2. The data table I constructed

is shown as Table 8 in Appendix B. In Table 8, zip codes with low alcohol availability are color coded light green, while zip codes of moderate availability are colored light yellow, and zip codes of high availability are colored light red.

- I obtained data on the fourteen leading causes of death for each county in VA from the Virginia Center for Health Statistics for the years of 2000-2002. A (truncated) sample of adjusted data is found in Table 9 in Appendix B.
- I obtained data on resident alcohol induced deaths by Race and Sex by underlying causes of death, as shown Table 10 in Appendix B.
- I obtained data on resident alcohol induced deaths for Fairfax County by Zip Code and Race/Sex, as shown Table 11 in Appendix B.
- I used the DMV Crash Facts database on a state and overall county. This information is broken down by exact street intersection. Every crash incident that had occurred in the Commonwealth of Virginia, and mentions if consumption of alcohol was involved in the incident. It also mentions if there was a crash, fatality, type of vehicle, individuals involved, their age, etc. is also in the DMV Crash Facts database. An example of this data is shown in Table 12 in Appendix B.
- I obtained data on suicide deaths by residential zip code in Fairfax County.
- I obtained motor vehicle injury deaths by residential zip code in Fairfax County.

- Crime data for each subcensus tracts and police patrol area were obtained. An example of this data is shown in Table 13 in Appendix B.
- I obtained race and ethnic origin for all 47 zip codes in Fairfax County. Precise count and percent of each race, ethnic origin, and gender for each zip code in Fairfax County.
- I obtained data on the population for each zip code broken down by age and gender.
- I obtained data on job class with race for Fairfax County.
- I obtained the following demographic data:
    - Age by language spoken at home by ability to speak English for the population five years and over for zip codes within Fairfax County from 2000 Census,
    - Household language by linguistic isolation for zip codes within Fairfax county from 2000 Census,
    - Household income in 1999 for zip codes within Fairfax County from the U.S. Census Bureau,
    - Median household income in 1999,
    - Poverty status in 1999 by age – for all zip codes within Fairfax County,
    - Ratio of income in 1999 to poverty level for all zip codes in Fairfax County,

- Poverty status in 1999 by age by household type for all zip codes in Fairfax County,
- Poverty status in 1999 of families by family type by presence of related children under 18 years by age of related children for all zip codes in Fairfax County,
- Imputation of veteran status for the population 18 years and over for all zip codes in Fairfax County,
- Imputation of period of military service for civilian veterans 18 years and over,
- Imputation of length of military service for civilian veterans 18 years and over,
- Imputation of earnings in 1999 for the population 16 years and over – percent of earnings imputed,
- Poverty status in 1999 of individuals not in families by imputation of individuals' income – percent of income imputed,
- Poverty status in 1999 of individuals in families by imputation of family income – percent of income imputed,
- Armed forces status by school enrollment by educational attainment by employment status for the population 16 to 19 years (white alone),
- Sex by employment status for the population 16 years and over (white alone),

- Sex by employment status for the population 16 years and over (black or African American alone),
- Sex by employment status for the population 16 years and over (American Indian and Alaska native alone),
- Sex by employment status for the population 16 years and over (Asian alone),
- Poverty status in 1999 by age (white alone),
- Poverty status in 1999 by age (black or African American alone),
- Poverty status in 1999 by age (American Indian and Alaskan Native alone),
- Poverty status in 1999 by age (Asian alone),
- Poverty status in 1999 by age (native Hawaiian and other Pacific Islander alone), and
- Poverty status in 1999 by age (some other race alone).

3.4 Impact on Community-Based Programs

This research carries a potential for high impact. Of course, I believe that the ability to simultaneously assess the impact of interventions on the probabilities of acute outcomes will allow for a choice of strategies that will reduce societal cost both in human terms (e.g. reduction of unnecessary deaths) and in financial terms (e.g. costs society incurs when prosecuting criminal activity related to undesirable alcohol behaviors). Moreover, the use of the relatively homogeneous clusters of agents in the model formulation has the added advantage of identifying at-risk subpopulations, and the dynamics of their adverse alcohol related behaviors. Indeed, the dynamic character of the simulation makes it possible to identify specific times, places, and circumstances for adverse behaviors for all the subpopulations. This makes subpopulation specific community-based programs possible. Alcohol abusers and alcoholics often need intense treatment therapies, including detox, educational programs, and medical treatment. In contrast, the user who is physically less tolerant of alcohol is often more profoundly affected by acute outcomes and usually needs a different type of treatment. Identifying the subpopulations at risk for both types of behaviors reduces overall societal cost by targeting appropriate treatment. Fairfax County specifically has community-based programs such as the Alcohol Safety Action Program, the INOVA Comprehensive Addiction Treatment Services, and a very large Alcoholics Anonymous community.

A final note on data availability is appropriate. Much of the data reported in this dissertation is publicly available. However, I cannot release some of the data, such as data related to court records, Virginia Department of Motor Vehicles, and health related data, that are available directly through government sources. The ftp site, ftp://www.galaxy.gmu.edu/pub/datasets/ReleasableAlcoholData/, contains data used in my dissertation that I am able to release.

Chapter 4: Estimating the Probabilities

The general strategy in estimating the probabilities was to use a frequentist approach based on the data I have collected. For the most part, the data were not collected according to a randomized designed experiment so the relative frequencies are somewhat problematic. The basic structure of the directed graph that I used in my simulation is given below. I begin by selecting a zip code. There are 47 zip codes and the selection of a zip code region is made proportional to the population within the zip code. I next choose an agent within the zip code. The agent is chosen based on the joint distribution of ethnicity and job class. The joint distribution was based on data from the Bureau of Labor Statistics. Unfortunately, data at the Fairfax County level are not available. The joint distribution is given in Table 1.

Table 1: Ethnicity versus job class joint probabilities

(Ethnicity x Job Class) Probabilities			
	White Collar	**Blue Collar**	**Unemployed**
White	0.337	0.613	0.052
Black	0.236	0.657	0.108
Hispanic	0.160	0.763	0.077

My next step is to decide whether or not the agent selected is a misuser of alcohol or not. “Misusers” are defined as individuals who are either alcohol abusers or alcohol dependent as defined in the NLAES data.

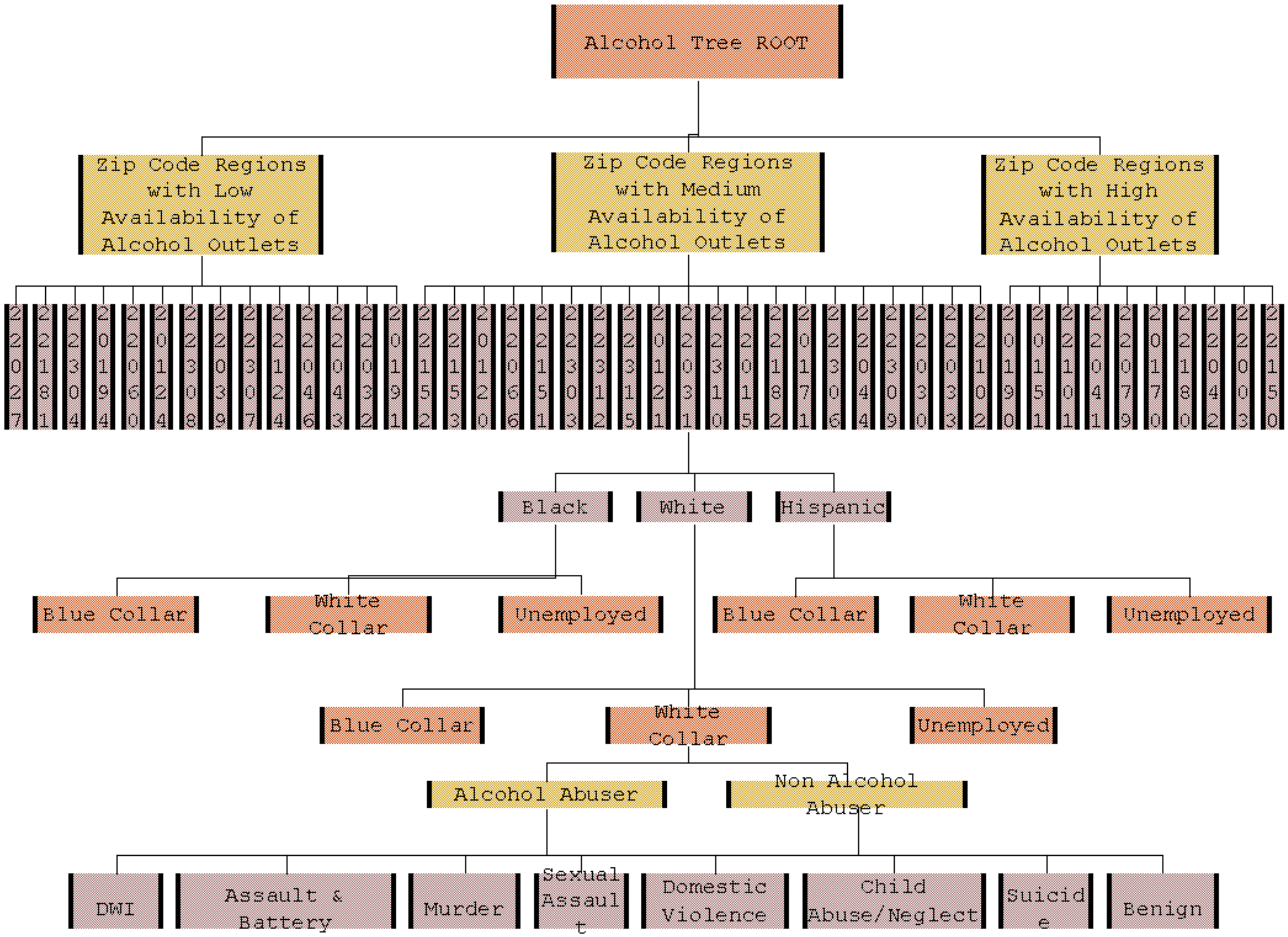


Figure 3: Alcohol Tree Directed Graph

The conditional probability of being a misuser is dependent on ethnicity, job class and zip code. The NLAES study provides the conditional probability of being an alcohol misuser conditioned on job class. The U.S. Bureau of Labor Statistics provides the joint distribution of ethnicity and job class. Finally the U.S. Census Bureau provides data on ethnicity by zip code. In order to calculate the conditional probability of being a misuser given ethnicity, job class and zip code, I made an assumption of conditional independence among the three probabilities mentioned above. Based on that assumption, I was able to calculate the desired conditional probability, which is given in Table 14 in Appendix B.

The probabilities given in Table 14 depend on ethnicity, job class, and zip code, but do not take into account the availability within the zip code. In order to approximate the availability effect I make some assumptions. Within a given zip code, let n_m be the number of misusers, n_n be the number of nonmisusers, n_p be the population of the zip code, and n_o be the number of outlets. I assume $n_o \leq .5n_p$. I further assume that n_m is proportional to n_o and that if $n_o = .5n_p$, then $n_m = n_p$ and if $n_o = 0$, then $n_m = 0$. Thus $n_m = 2n_o$. Thus discounting the ethnicity, job class, and zip code factors, the $n_n = n_p - n_m$. Let $P(m \mid e,j,z)$ be the probability of being a misuser given ethnicity, job class, and zip code and $P(m \mid e,j,z,a)$ be the probability of being a misuser given ethnicity, job class, zip code, and alcohol availability. The excess probability due to availability is calculated as follows

$$P(m \mid e,j,z,a) = P(m \mid e,j,z)(1 + 2n_o/n_p)$$

This is a working approximation that clearly needs to be calibrated by real data. The maximum value of this factor is 2. Thus it is possible to double the conditional probability of being a misuser depending on the availability.

I made the assumption that the probability of an acute outcome depends only on whether or not the agent is an alcohol misuser or not. This is not a realistic assumption in general. The first order justification of this assumption is that once someone is seriously under the influence of alcohol, the ethnicity, job class, and home location does not really matter. Realistically these factors and many others such as gender and age do matter. However, I did not have the data available to include all appropriate conditioning variable in this dissertation. As a crude first order approximation, I believe this assumption is adequate. The distribution of acute outcomes was calculated based on five years of police records within Fairfax County. For lack of better information I made a crude approximation that an alcohol misuser is five times as likely to have an acute outcome than a nonmisuser.

Chapter 5: Alcohol Tree Simulator

5.1 Overview

5.1.1 Technologies Used

- Java (J2SE 1.4 – original Alcohol Tree simulator)
- J#.NET (.NET Framework 1.1 – converted Alcohol Tree simulator).
- ASP.NET (.NET Framework 1.1, IIS 6.0 – web based front-end for the Alcohol Tree simulator library [written in J#.NET]).
- [D]HTML, Javascript, CSS (IE 5.0 or higher – client side application functionality).
- 3D Studio Max version 7.0
- Adobe Photoshop
- Viewlet Builder

5.1.2 Terminology

HTML is the coding language used to create Hypertext documents for use on the World Wide Web. HTML is a standard of the World Wide Web Consortium (W3C).

DHTML, or Dynamic HTML, is a method of combining HTML, CSS, DOM, and scripting languages [such as Javascript or ECMAScript] to allow for dynamic client-side manipulation of presentational components. When used appropriately, this can eliminate the need for a server request each time an action is to be performed, dramatically increasing the speed of interaction with the application.

CSS, or Cascading Style Sheets, are a specification for the presentation of HTML marked documents. Cascading Style Sheets work like a template, allowing Web developers to define styles for individual HTML page elements. CSS is a standard of the World Wide Web Consortium (W3C).

DOM, or the Document Object Model, is a programming interface that allows HTML pages and XML documents to be created and modified as if they were program objects. DOM makes the elements of these documents available to a program as data structures, and supplies methods that may be invoked to perform common operations upon the document's structure and data. DOM is both platform- and language-neutral and is a standard of the World Wide Web Consortium (W3C).

Javascript [formally known now as ECMAScript – as defined by the ECMA-262 standard] is a scripting language originally developed by Netscape. It is commonly used to make HTML documents more interactive, as it allows direct access to the underlying page DOM. Despite its name, JavaScript is not related to Java.

ASP.NET (sometimes referred to as ASP+) is the latest version of Microsoft's Active Server Pages technology (ASP). ASP.NET is drastically different than its predecessor in three major ways:

- It supports code written in compiled languages such as C++, C#, and J#.
- It features server controls that can separate code from the content, allowing WYSIWYG editing of pages (when using the Visual Studio .NET Interactive Development Environment [IDE]).
- It is fully Object-Oriented, and based on the .NET runtime – as such, it has full access to the underlying .NET class library.

Although ASP.NET is not backwards compatible with ASP, it is able to run side by side with ASP applications.

A DLL, or Dynamic Link Library, is a file of functions – compiled, linked, and saved separately from the processes that use them. Functions in DLL's can be used by more than one running process. The operating system maps the DLL's into the process's address space when the process is started up or while it is running.

WYSIWYG is an acronym for “what you see is what you get.” WYSIWYG HTML Editors like Dreamweaver or Frontpage let one create web pages by displaying exactly how it will look in a browser. Because of this, intrinsic knowledge of HTML is not necessary – the use of WYSIWYG editors is problematic, however, because of their use of non-standard, proprietary and deprecated mark-up. Therefore, I did not use any WYSIWYG editors during the course of this research.

Java – Developed by Sun Microsystems, Java is a network-oriented programming language that is specifically designed for writing programs that can be safely downloaded to the computer through the Internet and immediately run without fear of viruses or other harm to the computer or files. Using small Java programs (called “Applets”), Web pages can include functions such as animations, calculators, and other fancy tricks. Java is a simple, robust, object-oriented, platform-independent multi-threaded, dynamic general-purpose programming environment. It is best for creating applets and applications for the Internet, intranets and any other complex, distributed network.

J#.NET is a powerful tool for Java-language developers who want to build applications and services on the Microsoft .NET Framework. It targets the .NET Framework version 1.1, is fully integrated with Visual Studio .NET, and provides added support for building Mobile Web applications. J#.NET includes technology that enables users to migrate Java-language programs to the .NET Framework [in minimum time]. Existing applications developed with Java can be easily modified to execute on the .NET Framework, interoperate with other Microsoft .NET-connected languages and applications, and incorporate .NET functionality such as ASP.NET, ADO.NET, and Windows Forms. It should be noted that J#.NET is not a tool for developing applications intended to run on a [Sun] Java virtual machine. Applications and services built and compiled as J# [.NET] code will run only in the .NET Framework; they will not run on any Java virtual machine. J#.NET is an independent development by Microsoft – It is neither officially endorsed nor approved by Sun Microsystems, Inc.

“Just-In-Time” or JIT refers to a compiler for the Java language that allows interpreted Java programs to be automatically compiled into native machine language on the fly, for faster performance of the program.

5.1.3 Overview Description

The simulation components of this research are written as a mix of Java and J#.NET. These are both Object-Oriented development languages implementing the Sun Java Language Specification. Java is the official Sun implementation of this spec – whereas J#.NET is a Microsoft derivative utilizing Microsoft's .NET Framework library. These two languages are not binary compatible; however they are mostly source-compatible.

These different Java flavors are both compiled languages, similar to C or C++. The major difference is in the “level of compilation” achieved. In C, for instance, the code would be compiled down to raw x86 (Assembly language) instructions, which would form the binary executable. One problem with this method of compilation is portability – if code is run on a different platform or architecture, many system calls and hooks using #IFDEF preprocessor logic must generally be changed. Basically, the compiler itself is allowed to conditionally include portions of code depending on architecture type – for instance to target a WIN32 architecture the code is placed inside an #IFDEF WIN32 block, in which case it would *only* be compiled if the architecture matched. This is standard C and C++ fare; in Java, however, since the binary executable contains Virtual Machine (VM) code [one level above Assembly Language], the VM can execute the code in the same fashion regardless of which system architecture is currently being used. One could think of this, in a sense, as a form of integrated compatibility layer built into the language itself. The Virtual Machine, however, contains a very sophisticated method of dynamically compiling code “Just-In-Time” (JIT) as the executable is running. One additional benefit of this JIT Compilation is that it can perform architecture-specific optimizations at runtime, something a traditional C or C++ program could never [reasonably] achieve. These optimizations can dramatically increase the speed of the executing Java application – sometimes making them even faster than their C and C++ counterparts.

This application is currently hosted on an IIS6.0 web server running ASP.NET with the .NET Framework version 1.1. The web host is uplinkearth.com.

5.2 Description

The actual simulation aspects of this research are functionally decoupled from the [presentational] front-end, and are packaged as a separate library [DLL]. Throughout development, efforts were made to utilize an object-oriented methodology wherever applicable. Fundamentally, the research can be broken down into three main processes of functionality:

i. Alcohol Tree simulation process.

ii. The map generation process.

iii. The presentation and [client-side] user-interaction process.

5.2.1 Part 1

The Alcohol Tree simulation process is the one that actually performs the simulation given the number of agents and runs. This program is composed of eight Java classes. Each Java class is contained in a file of the same name; e.g. a class named “Node” will be in a file named “Node.java”. The program is designed using a “tree”. Trees in computing are similar to real-life trees. A computing tree has a “root” node along with one or more hierarchal child nodes. The root node is the node at the top of the tree, with no parent in the hierarchy.

The “AlcoholTree” class is the main class for the program. It performs the actual simulation given the number of agents and output file it should use. There is a command line version of the simulation using:

java AlcoholTree 1000000 C:\fairfax.txt

(In the same command line directory as the compiled java code [.class files]).

The “Node” class is the base class for the program. It describes a default node capable of having multiple children [defined by each individual class]. The “OutletNode” class extends the “Node” class (i.e. [that is] it intrinsically inherits the attributes of the “Node” class). Each of the other classes, in turn, inherits from another class. The order of inheritance is shown below:

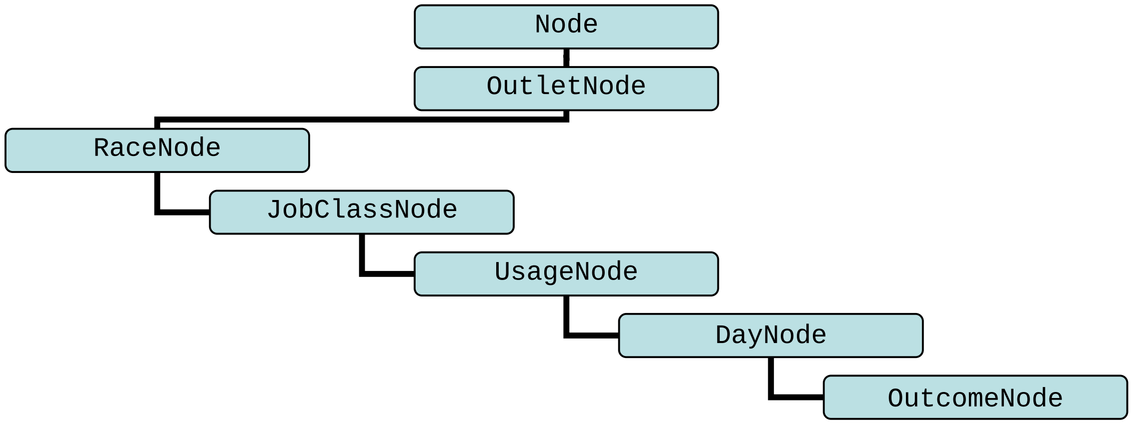


Figure 4: JAVA class nodes and their relationship

In reality, there is no reason for each of the different node classes to inherit from the previous node class. They could all just as easily inherit from the base “Node” class itself. The only reason for such a structure is if the subsequent child node classes require access to variables or functions in the parent node class. This is not the case with this alcohol program.

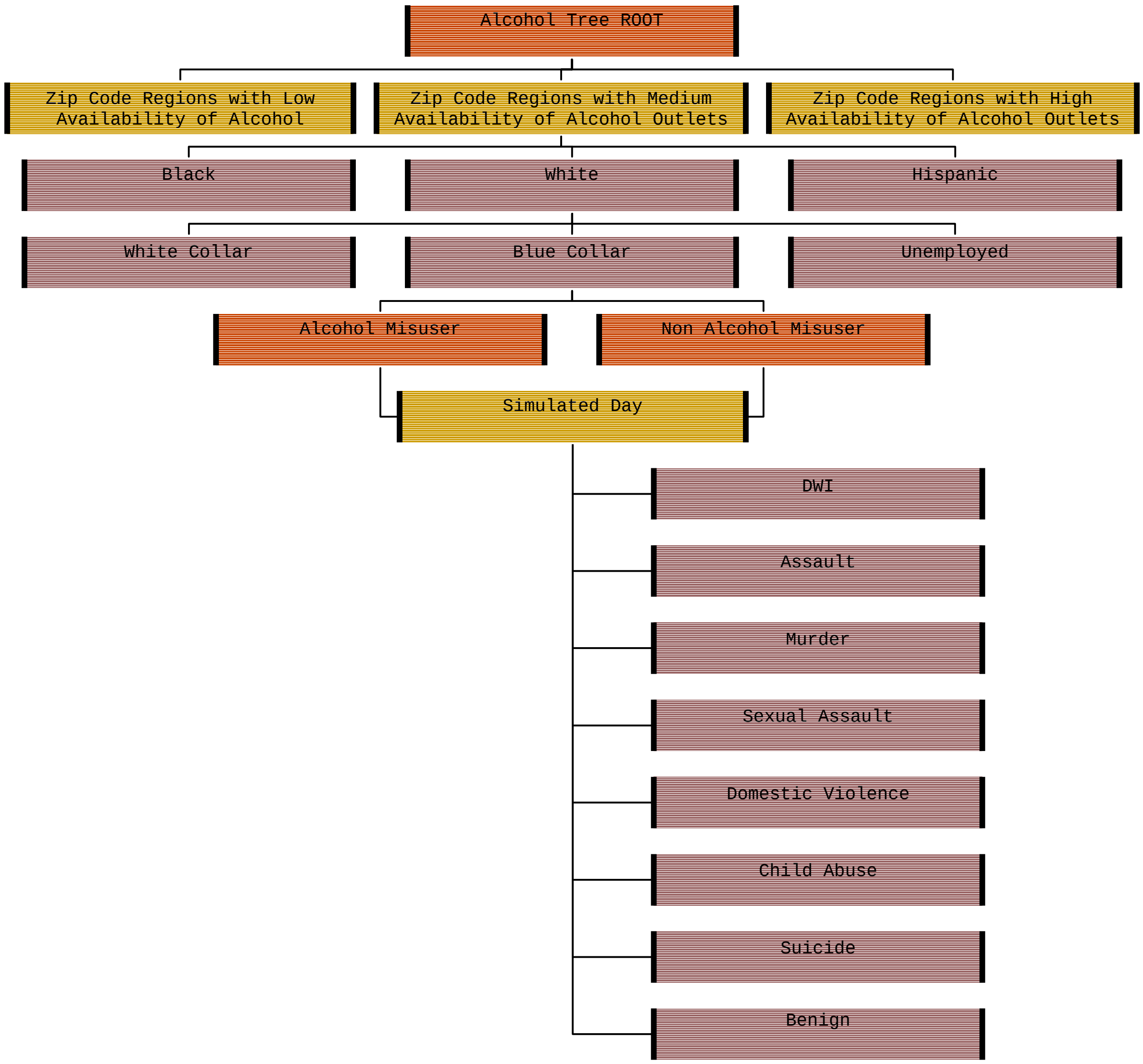


Figure 5: Abbreviated alcohol tree. The bottom portion is replicated for each of the nine ethnicity – job class pairs.

Note: The above tree does not include every possible combination due to size constraints. Each level of the tree is repeated for each node above it. For example, the “Non Alcohol Misuser” node also has a “Simulated Day” child node with corresponding acute outcomes (seven of them). Similarly, there are “Alcohol Misuser” and “Non Alcohol Misuser” child nodes for the “White Collar” and “Unemployed” nodes. Interpreted literally, it may appear that only the “Blue Collar” node has “Alcohol Misuser” and “Non Alcohol Misuser” nodes, which is not the case.

Each of the Node classes contains probability values that are utilized by the simulation in determining whether an acute outcome occurs. The probabilities correspond to the different types of the node (e.g. – there are three types of OutletNodes – one for each low alcohol availability, medium alcohol availability, and high alcohol availability).

5.2.2 Part 2

The map generation process is where I actually begin to make sense of the simulation results from the previous stage.

The first thing, which occurs at this point is the processing of the actual [Fairfax County] GIS shapefile data. This occurs in two parts;

- The processing of the .SHP file – which contains zip code regional point data as paired latitudinal and longitudinal coordinates.
- The processing of the .DBF file – which associates an actual zip code number with the grouped coordinate data previously read from the .SHP

file. This .DBF file also contains additional information, such as the computed regional area.

The .SHP, or ESRI Shapefile, data format stores non-topological geometry and attribute information for the spatial features in a data set in a binary file. The geometry for a feature is stored as a polygonal “shape” comprised of a set of vector coordinates (representing latitude and longitude).

An ESRI “shapefile” consists of a main file, and a dBASE table (of the same name, with .DBF extension). The main file is of variable-record-length; each individual record describing a shape with a list of its vertices. The dBASE table contains feature attributes with one record per feature, causing a one-to-one relationship between the geometrical and attribute data between files.

For reading in the ESRI shapefiles, I am using slightly modified classes [components] originally from the CCmap application. What this is basically doing behind the scenes is opening and reading in the raw binary data from the underlying file “stream,” according to the ESRI data specification. My program then stores these data locally for later use.

From here, I “scale” the coordinate data from latitudinal and longitudinal coordinates to X and Y pixel values. To do this, I use this simple operation, given:

x = Original X value [Longitude]

bL = Minimum X value in the entire GIS shapefile [bounds].

bH = Maximum X value in the entire GIS shapefile [bounds].

d = Resulting image dimensions [width and height].

$$\text{Scaled X} = \frac{(x - bL) \times d}{bH - bL}$$

And, likewise:

$$\text{Scaled Y} = \frac{(y - bL) \times d}{bH - bL}$$

At this point, I have an internal list of zip code regions with their scaled X and Y pixel positions ready to be drawn to the screen. Each of these is logically organized as a “ZipcodeRegion” object.

The next step is calculating the maximum number of acute outcomes in a single zip code region. I need to use this maximum value later on, to dynamically adjust the intensity coloring.

Now I actually begin to outline and “fill” the actual zip code regions on the map:

- The zip code regions are simply drawn to the screen as closed polygons, using their scaled X and Y values.
- The zip code region intensity is then calculated, to determine the color used to fill the particular region. This is calculated as the percent of acute outcomes occurring in an individual region relative to the maximum

number of acute outcomes happening in any region [calculated previously]. Based on this “intensity value”, the actual color value [amount of red, green, and blue] is scaled accordingly. A computer screen is comprised of approximately one million pixels. Each of these pixels has a red, green, and blue component to it – a value of 0 (no color) to 255 (full color) is assigned to each of these components to determine its color. Now, white is created by combining all of the colors, so likewise, white would be represented as { R = 255, G = 255, B = 255 }. To vary the shading between white and red, I modify the green and blue values (red is kept constant at 255) based on the intensity value previously calculated. The default value for each polygon is white (R, G, B = 255). I decrease the B and G values inversely proportional to the ratio of the outcomes in one area to the maximum value in any area. Because I am using the intensity value to directly determine the amount of green and blue of the individual pixel, I obtain a maximum of 254 shades between red and white. A solid red represents a region with the highest number of acute outcomes, while a solid white region represents a region with the lowest number of acute outcomes.

I display pie charts to represent Black, White, and Hispanic populations within each individual zip code region. These are simple percentage calculations, using our population data, which is used to determine circular angles for each of the three chunks of the pie chart. The charts are then drawn at the region centroids. The color representation for the pie slices is White => White race, Black => Black (African American) race, Tan => Hispanic ethnicity.

5.2.3 Part 3

The presentation layers of this research are two-fold; the server back-end component, an ASP.NET script, and the client, the users' web browser.

I make use of internal sessions [via cookies] to maintain persistence across web-requests, as HTTP by nature is a stateless protocol.

By default, a single web page “hit” will generate two actual underlying requests:

- The first will output the Alcohol Tree Simulator interface, with its buttons and fields.
- The second will retrieve the actual map image [created previously] to be displayed.

To begin, I call the library [from Part 1] to perform the actual simulation, given the number of agents to be simulated. This simulation only occurs on the first of these two underlying requests, as it would be erroneous to re-simulate before the map image is generated and shown.

At this point, the interface is presented to the user, and allows client-side interaction [via “DHTML”, or Javascript] of region centroid, region name, alcohol establishment, and ABC store information]. The user is also presented with the underlying probability values used to generate the simulation results, with the ability to change and re-run the simulation with the new values.

The ABC and Alcohol Establishment store data are outputted by the ASP.NET script as [by default – hidden] HTML elements, which are simply [small] square colored boxes. These boxes are shown and hidden via DHTML and underlying DOM element access. More specifically, JavaScript is used to access the DOM – through which I access the Document's CSS elements, and modify accordingly [depending on whether I would like the elements to be hidden or visible].

To retrieve the latitudinal and longitudinal coordinates of the ABC stores and alcohol establishments, a separate tool was written to automate the [otherwise quite lengthy and error-prone] process.

This tool reads in the individual establishment locations from an XLS spreadsheet file and passed the address information on to a Google Maps query (http://maps.google.com/). The resulting page HTML was then analyzed and parsed to determine the actual latitude and longitude of the specified address. (This was determined to be hidden behind the scenes of the Google Maps interface, in a custom HTML tag, which defined these values as xPos and yPos attributes). The tool then stored these latitudinal and longitudinal values to a simple text file (one entry per line) to later be read and processed as necessary. Performing these lookups would have taken on the order of days to weeks, as there were literally thousands of addresses, which needed to be looked up. Using this automated method, it took one to two hours to both go through, and double check each of the address locations.

The map presented to the user is made interactive through the use of an HTML "imagemap," which causes an inline frame to display region-specific demographic information (static HTML files, auto-generated by a web-extractor tool for this purpose).

The user is presented with an option of displaying Simulation Statistics or Detailed Simulation Output. The latter will be described below; however it should be noted that the Simulation Statistic [percentage] values are calculated relative to the total number of *agents*.

Four tables are generated and presented to the user. The first shows the number of Acute and Benign outcomes for all Alcohol Misusers and Non Misusers. The second table shows the number of Acute and Benign outcomes for all White, Black (African American), and Hispanic individuals. The third table shows the number of Acute and Benign outcomes for all White Collar, Blue Collar, and Unemployed individuals.

The fourth table relates actual Fairfax county police record data to the simulation results, by Outcome type and calculates the standard deviation. The actual police record data are for a five-year period, so the data are rescaled to a one-year period. The simulation results are also scaled accordingly if one million agents were not used so that the standard deviation calculations are appropriate. Because Fairfax County has a population of approximately one million, I have settled on a nominal standard of one million agents. Although the simulation can be done with virtually any number of agents, I perform the scaling to make the simulated statistics consistent.

Alcohol Tree modified probability utilization:

A map key is generated using the maximum number of acute outcomes for a given region [I calculated this in Part 2] and dividing that into 12 color shades of varying intensity, between solid red and solid white. Each of the color values has an upper and lower bound to it, which is why the key appears “seamless”. The upper bound color is the one representing the high-end number of acute outcomes represented by the color shade. Likewise for the lower bound color. Therefore, if a key item had the boundaries of (0 to 50) and was colored from white at the bottom to slight pink on the top of the item boundary; a value of 25 would be one between these two points in color.

Color Coded establishment types:

To color code the establishment types, I again used the tool [previously described] to automate the retrieval of latitudinal and longitudinal coordinates from the Google Maps service. When the tool was originally used, and the data were stored to the initial text file (containing addresses, zip codes, and latitudinal/longitudinal coordinates)

The tool was modified to include alcohol license type in the resulting text file, representing the various licenses as integer values internally (zero if the license information cannot be determined, one if the establishment has a license to server alcohol on premise, two if the establishment has a license to server alcohol off premise, and three if the establishment has a license to serve alcohol both on and off premise).

In the ASP.NET script, after reading in the latitudinal and longitudinal coordinate data, I determine which color the establishment point should be based on the license value integer previously described. If the establishment has a value of two (off premise only) then it is colored blue, otherwise (on premise only, or on and off premise) it is colored green.

5.3 Additional Information

It should be noted that there are occurrences in the generated map where a single zip code has multiple physical boundary regions for it. The reason for this is that in the underlying GIS data I am using to generate the map, there are times when a single zip code contains multiple physically decoupled "polygonal regions," which in some cases are located within a different zip code region entirely. I do not attempt to determine these occurrences and remove them, as doing so would not be accurate because they do exist as a portion [albeit – decoupled] of a region. The GIS data I am using is the "5-digit Zip Code Boundary" Census data, which can be retrieved from: http://mason.gmu.edu/~riyengar/Fairfax/Fairfax.htm. The application is targeted for Internet Explorer Version 5.0 or higher.

The detailed simulation output is written for a detailed representation of the entire simulated "alcohol tree," displayed in its hierarchical form as child "nodes" of one another. At the root hierarchy are the three alcohol availability outlets [low, medium, and high]. Within each of these three outlets are race classes [black, white, and Hispanic]. Each of these race classes contains three job classes [white collar, blue collar, and unemployed]. Underneath the job classes lies node representing alcohol usage [misuser, or non misuser]. Within each of these two individual nodes lie eight final outcome nodes – one each to represent DWI, assault, murder, sexual assault, domestic violence, child abuse, and suicide, along with a final node to represent all simulated benign outcomes.

Regarding the final data table:

“Actual Incidents” is calculated as the total number of each incident type from the police data file divided by the number of years of data contained in the file [3].

Simulated Results are calculated as the total number of each individual outcome type [obtained by iterating over all alcohol tree node elements and retrieving ones, which apply [in this case, only outcome type nodes] multiplied by a “scaling” factor.

This “scaling” factor is only applied if the simulation is not run with one million agents. By scaling the total number of acute outcomes according to the number of agents the simulation was run for [based on the percentage of agents out of one million agents], the data between columns (Actual, and Simulated) becomes comparable.

To calculate the “Mean Simulated Results”, I wrote a small application. The [averaged] data are then be scaled accordingly [multiplied by ten] to more accurately represent the actual 1,000,000 “agents” desired.

Table 2: Mean simulated results

Average RFA	0.001365
Average RFB	0.948489
Average Number of Acute Outcomes	1365
Average Number of Benign Outcomes	948489
Average Number of Agents	949855
Average Number of Runs	1000000
Average Total DWI	707.5
Average Total Assault	132.2
Average Total Murder	7.4
Average Total Sexual Assault	33.9
Average Total Domestic Violence	168.2
Average Total Child Abuse	216.4
Average Total Suicide	99.8

These values are very close to the actual data [collected from the police file]. The data [rounded to whole numbers] are shown side by side below, for comparison:

Table 3: Comparison of simulated with actual acute outcomes

Outcome Type	Data From Police File	Mean Simulated Data
DWI	722	708
Assault and Battery	133	132
Murder	6	7
Sexual Assault	32	34
**Domestic Violence*	41	168
Child Abuse/Neglect	84	216
Suicide	49	100
Benign	998933	998635

*In addition to this Domestic Violence count, there is an [actual] Domestic Dispute count of 6270 disputes per year.

Based on 100 Monte Carlo replications, the Mean Square Error (MSE) =

$\frac{1}{100}\sum_{i=1}^{100}(\text{simulated}_i - \text{actual})^2$, where *i* is the varying acute outcome (i.e. DWI, assault and battery, murder, sexual assault, domestic violence, child abuse, suicide, and benign outcomes). The Absolute Deviation =

$\frac{1}{100}\sum_{i=1}^{100}\left|\text{simulated}_i - \text{actual}\right|$.

Chapter 6: Navigating the Website

Navigate to http://www.alcoholecology.com/

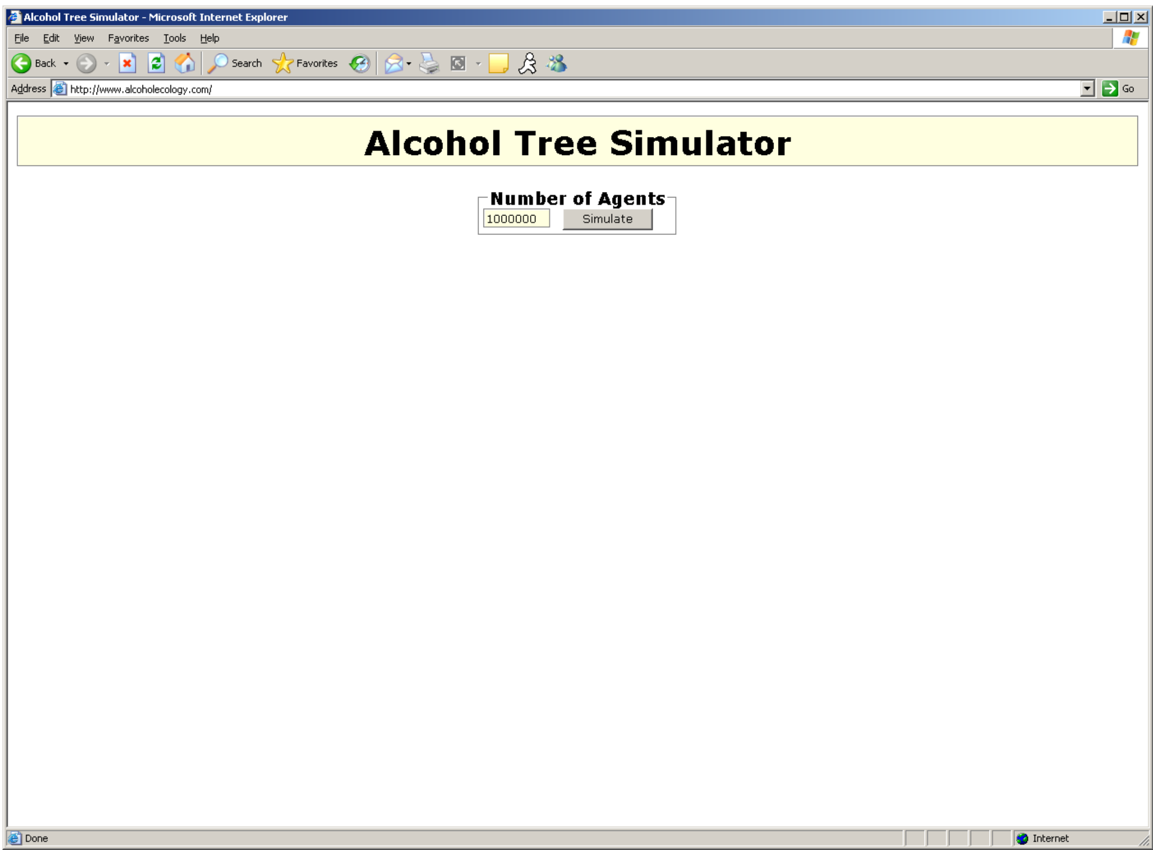


Figure 6: Opening screen of web simulator application.

The user is prompted to enter the number of agents. The default is 866295, which was the 2000 census population. An agent represents a single individual within Fairfax County.

The user runs the simulation by clicking on the 

button.

The user is presented with the simulation results and an accompanying color-shaded map of Fairfax County:

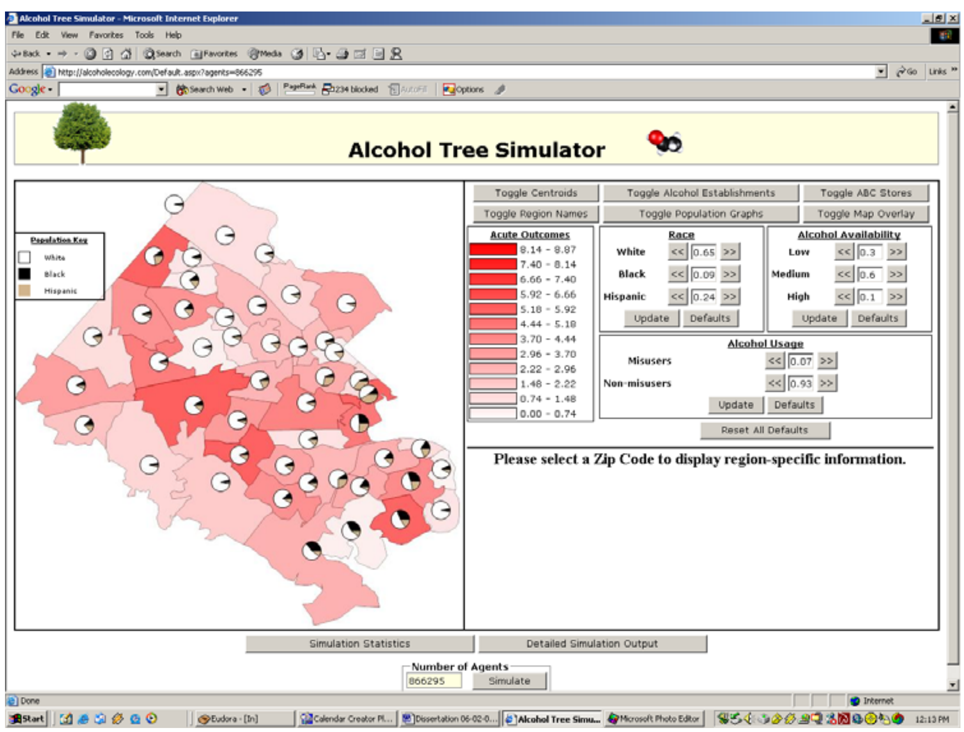


Figure 7: Screen shot after the simulation has been executed.

The map of Fairfax County is color coded by zip code regions. By default, each zip code region is overlaid with a pie chart indicating the White, Black, and Hispanic populations within that region. These are indicated by the colors white, black, and tan respectively. This is seen below:

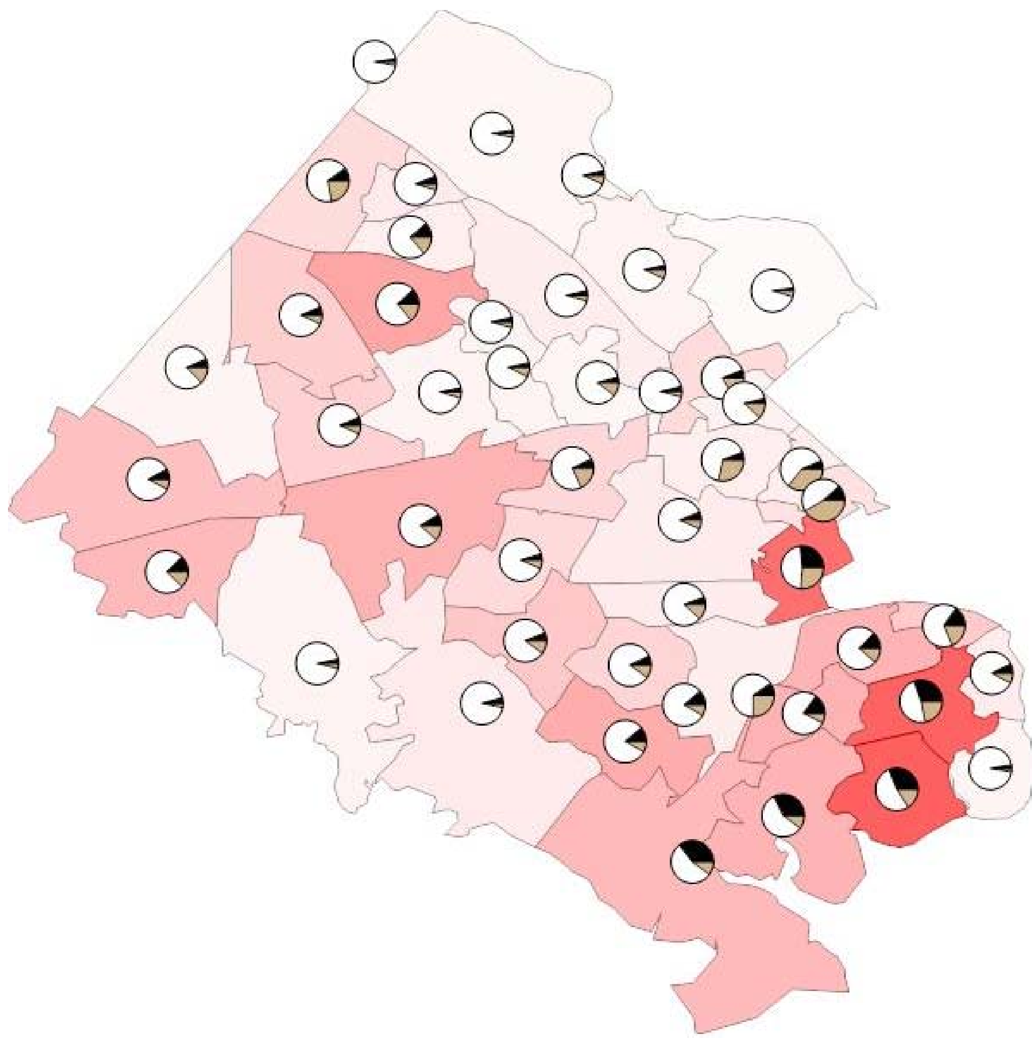


Figure 8: Fairfax County map with population pie charts. The pie charts indicate ethnic characteristics. White represents proportion of the White population, black the African American population, and tan the Hispanic population.

To the immediate right of the Fairfax map, there is a gradient "key" indicating the number of acute outcomes associated with each color value. In this example, for instance, the darkest shade of red would indicate 14 acute outcomes within a particular zip code region.

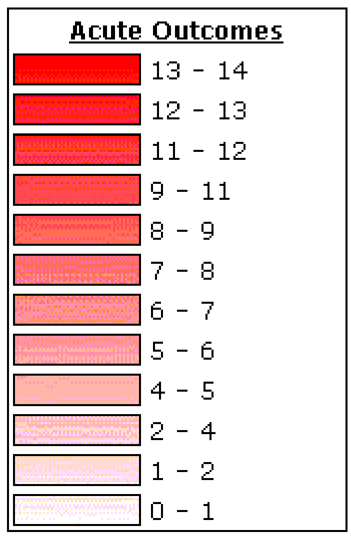


Figure 9: The color scale representing the number of acute outcomes.

Above this color key, there are six buttons:

The button actions are as follows in order of use:

Toggle Population Graphs: Disables/enables the display of the population pie charts overlaid on the zip code regions. For example, when disabled:

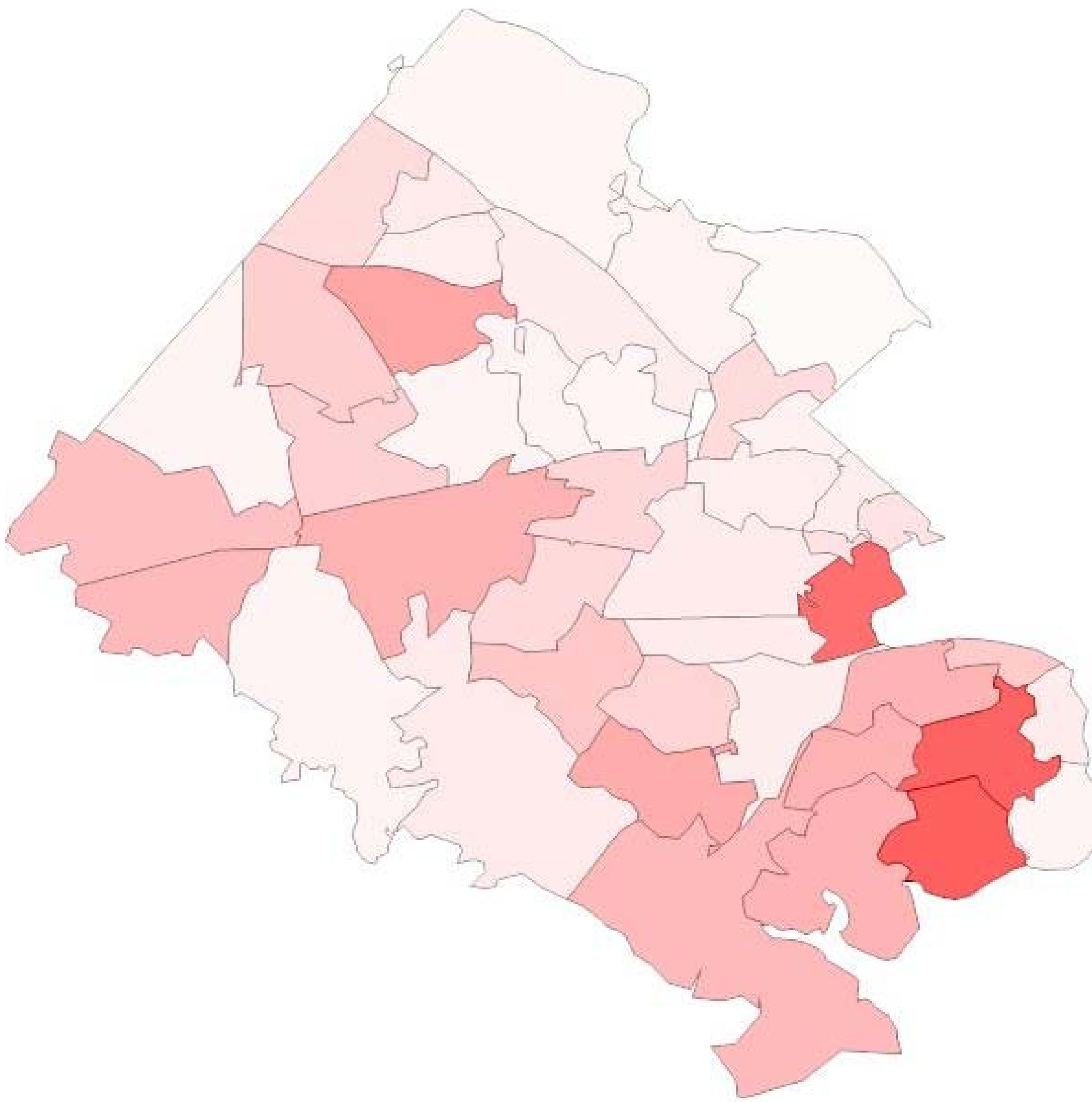


Figure 10: Fairfax County map with more saturated colors indicating higher incidence of acute outcomes.

Toggle Centroids: Enables/disables the display of markers at each zip code regions "centroid" (center of mass), as seen below (with population graphs disabled for visibility):

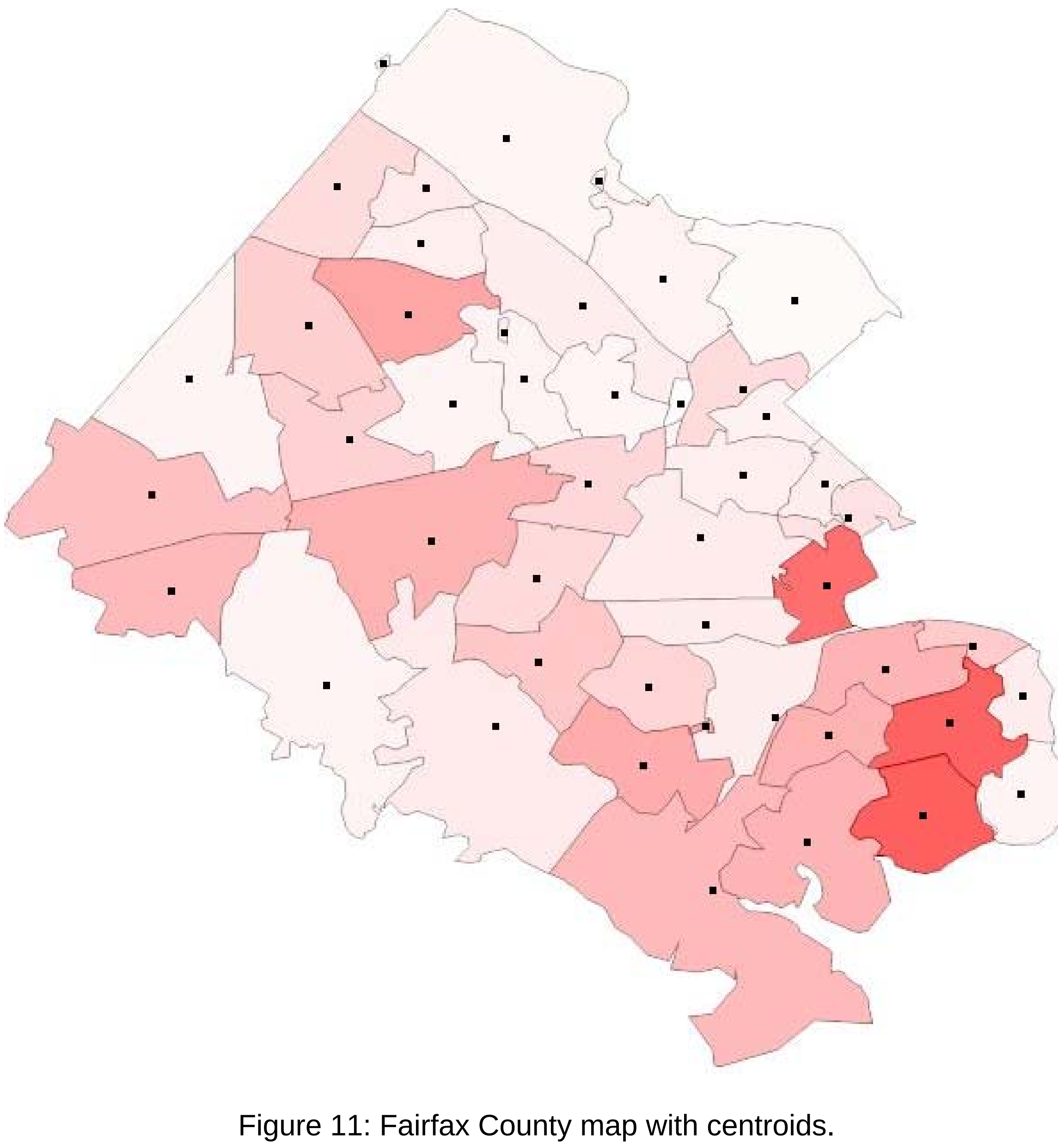


Figure 11: Fairfax County map with centroids.

Toggle Alcoholic Establishments: Enables/disables the display of [actual] alcohol establishments in Fairfax. These establishments are placed on the map in their physical locations, determined by their latitudinal and longitudinal coordinates:

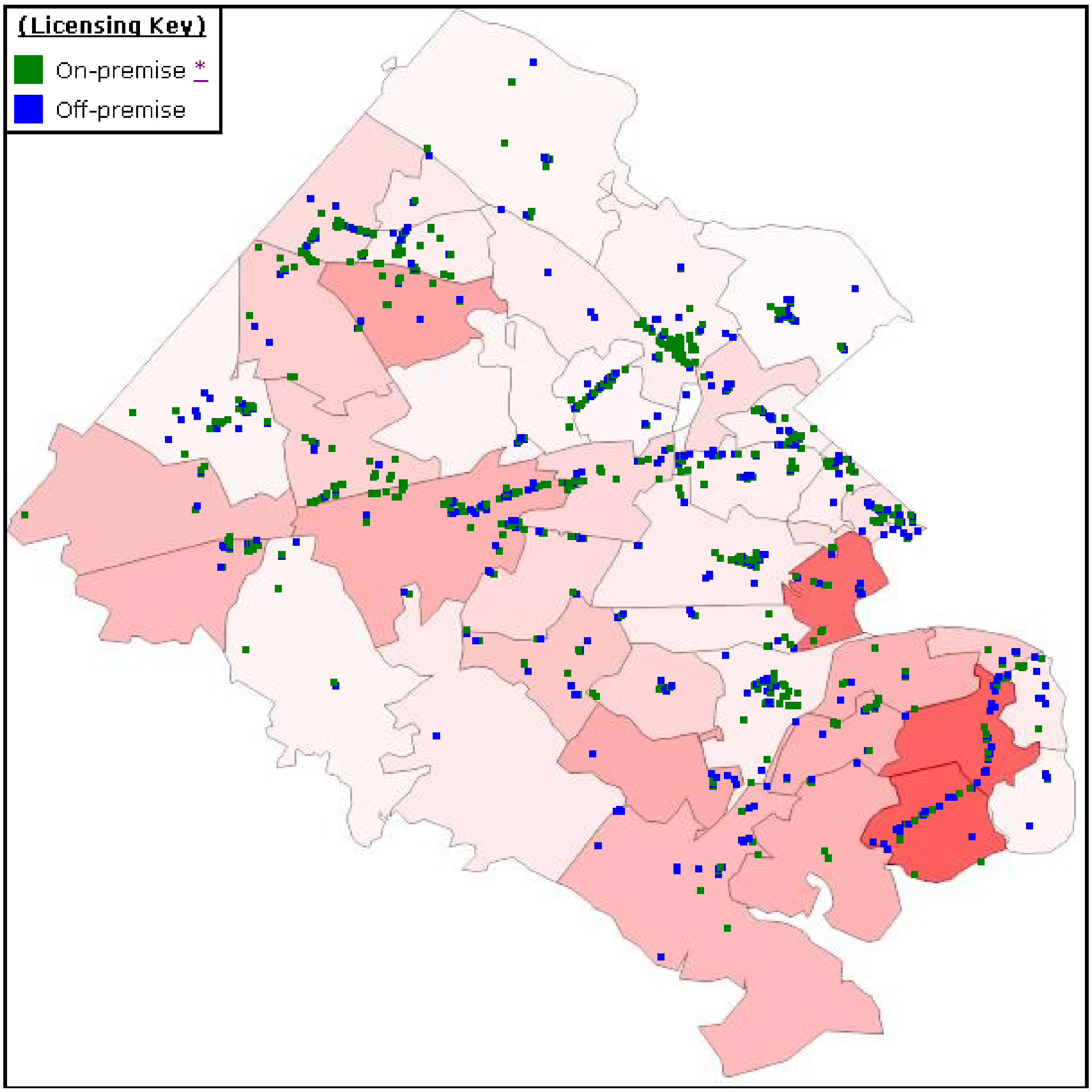


Figure 12: Fairfax County map with establishments licensed to sell alcohol

Toggle ABC Stores: Enables/disables the display of the Alcoholic Beverage Control (ABC) stores within Fairfax County. These stores are also placed on the map in their actual geographic locations. In the example below, the ABC stores are red, and displayed with the alcohol establishments enabled as well:

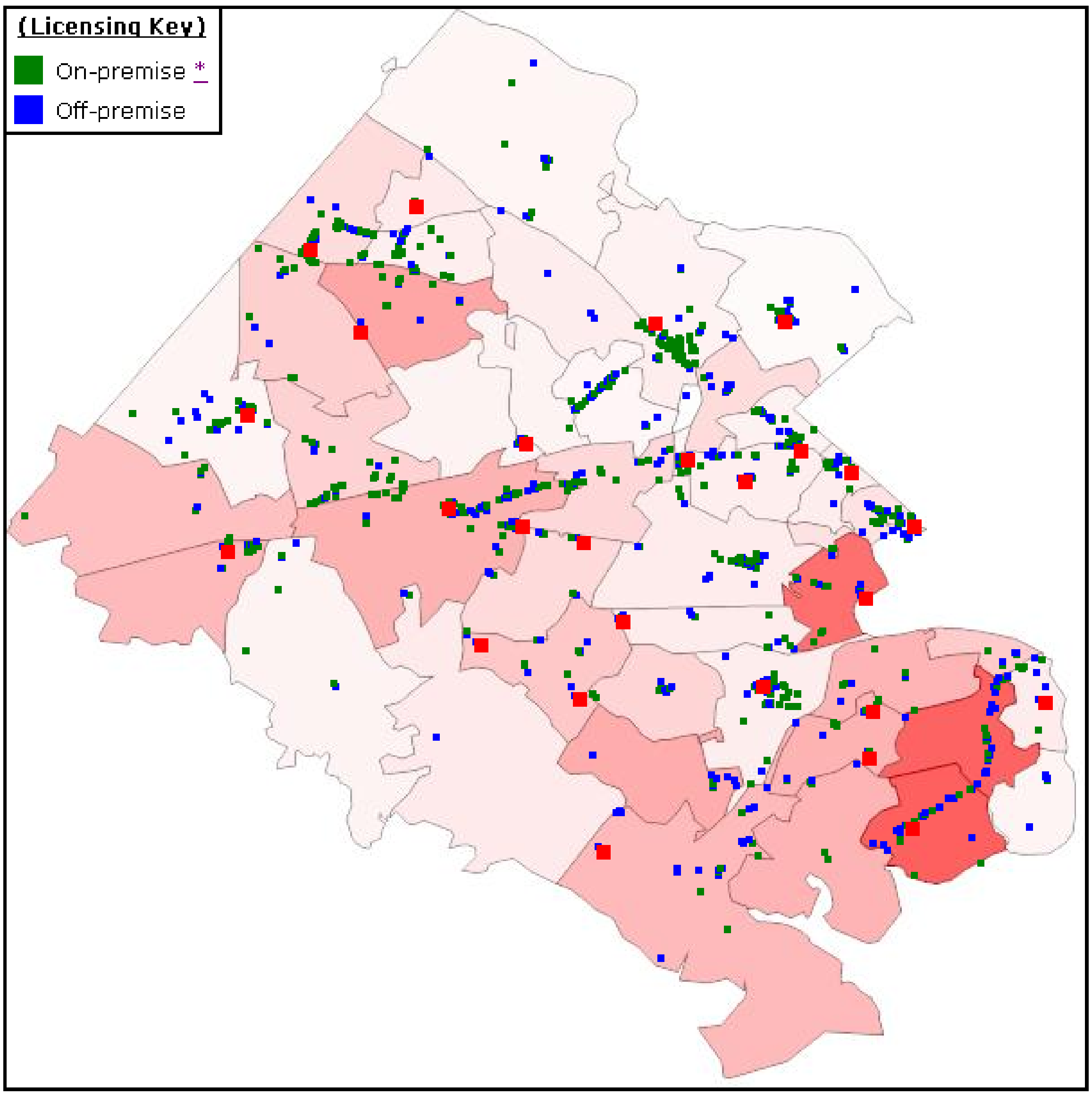


Figure 13: Fairfax County map with establishments licensed to sell alcohol and including the state-owned ABC stores.

Upon selecting the **Toggle Alcohol Establishments** button once more, they are removed from view – leaving just the ABC Stores and their red indicators:

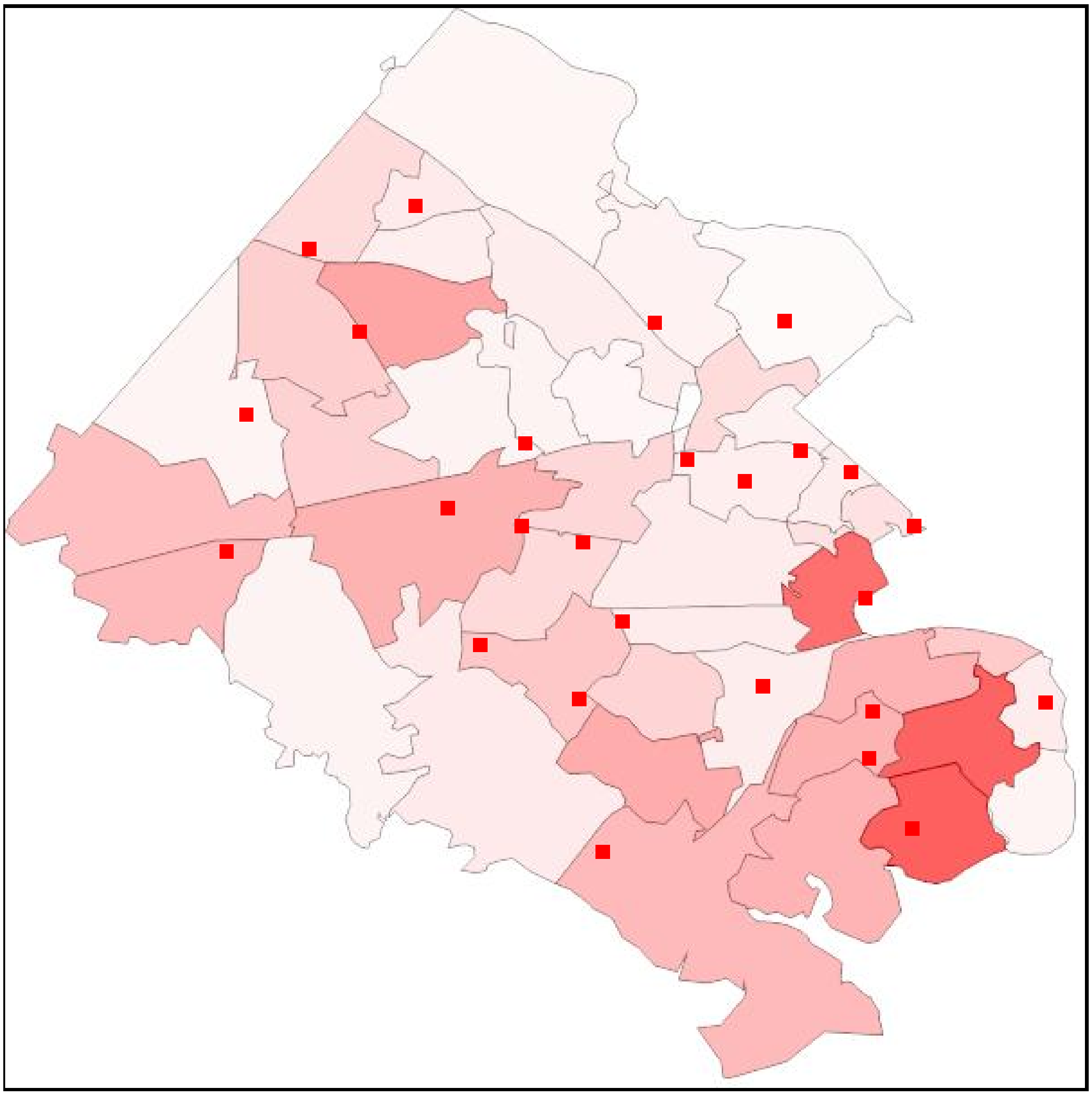


Figure 14: Fairfax County map with state owned ABC stores.

Toggle Region Names: Overlays each zip code region's "place" names:

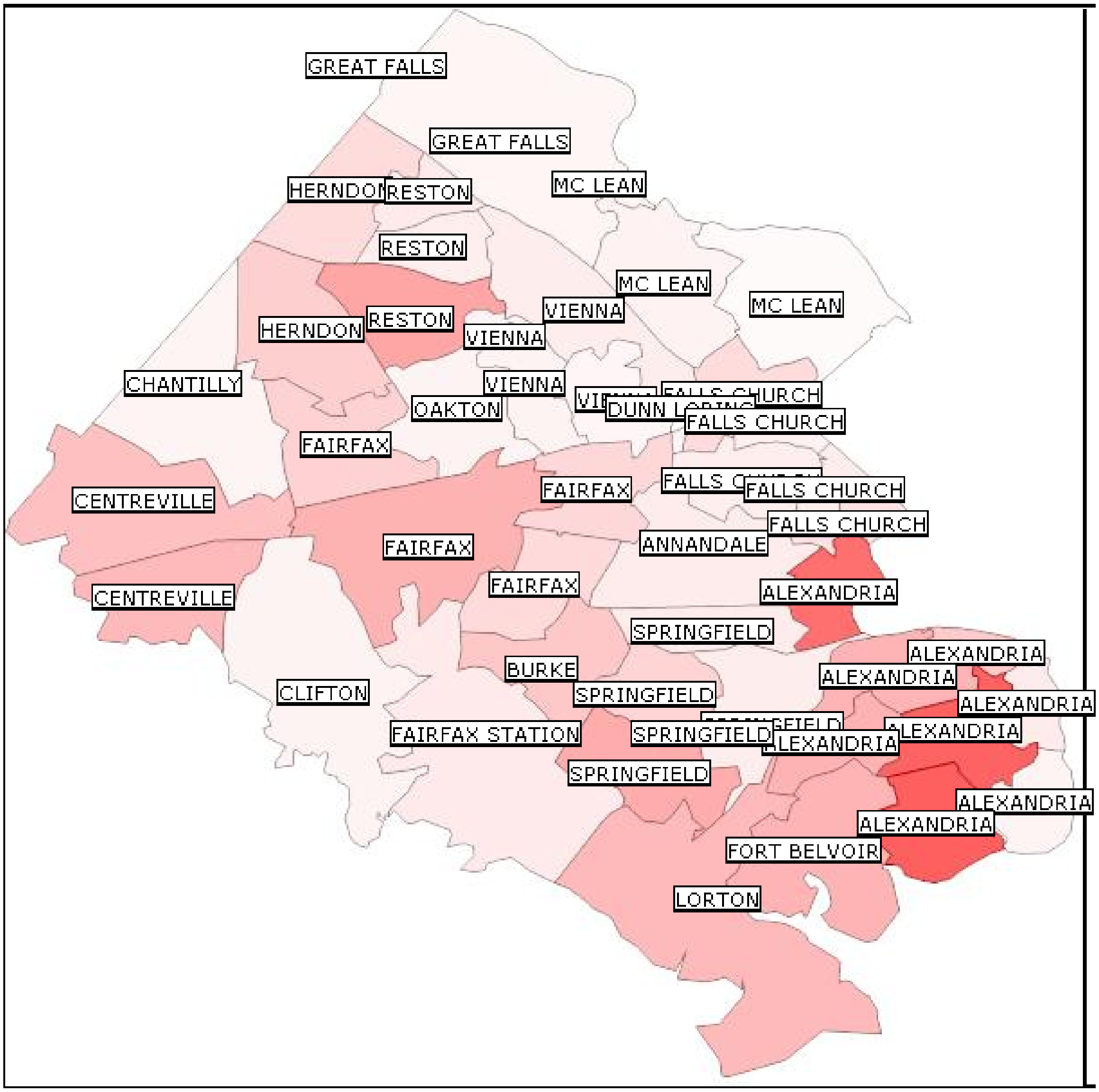


Figure 15: Fairfax County Map with zip code place names.

Toggle Map Overlay: This button toggles the overlay of the actual Fairfax County region map, including road names and zip codes. The actual map is overlaid with a 70% opacity level, so that the original map is slightly visible underneath it. It is interesting to note the subtle differences between the zip code region boundaries based on GIS data (underlying image) and police data (overlaid map showing roads). This is illustrated below:

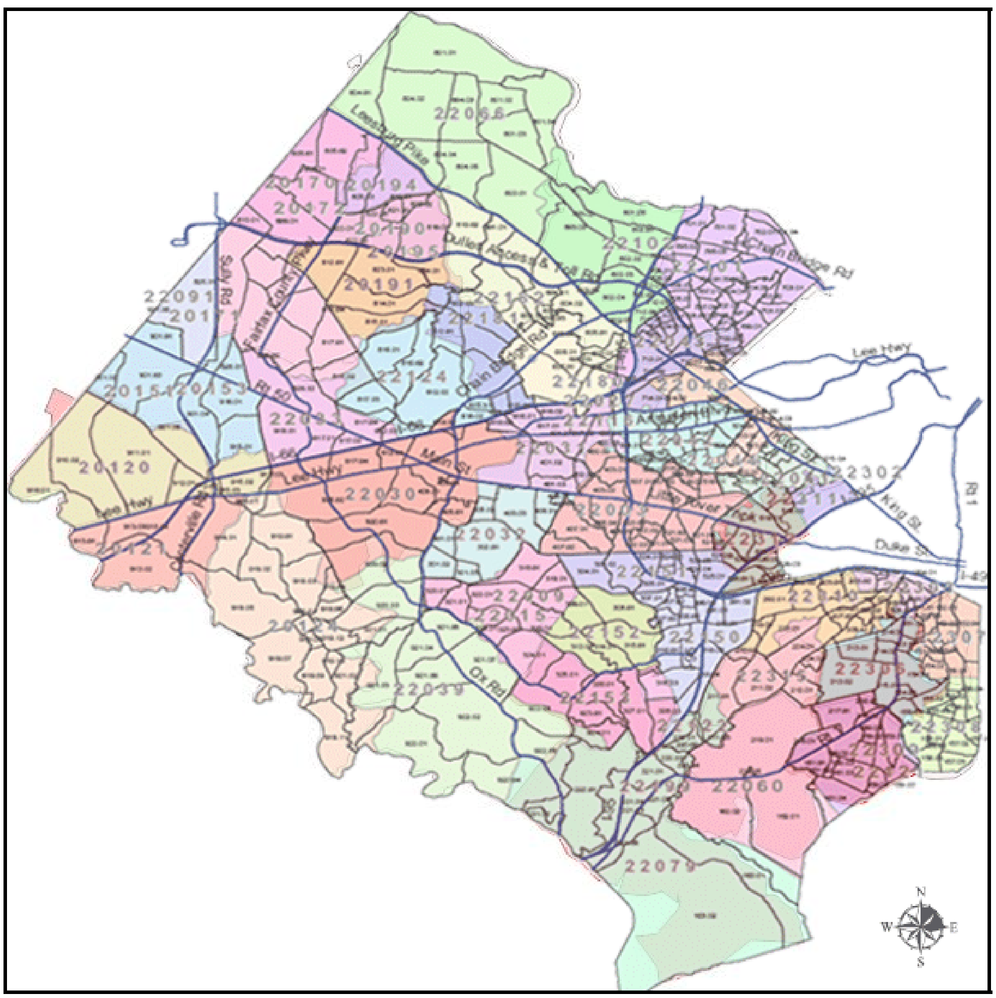


Figure 16: Fairfax County map with street overlay.

The user may also “mix and match” various button states. For instance, if the user enables map overlays as well as the display of alcohol establishments and ABC stores, the image below is generated:

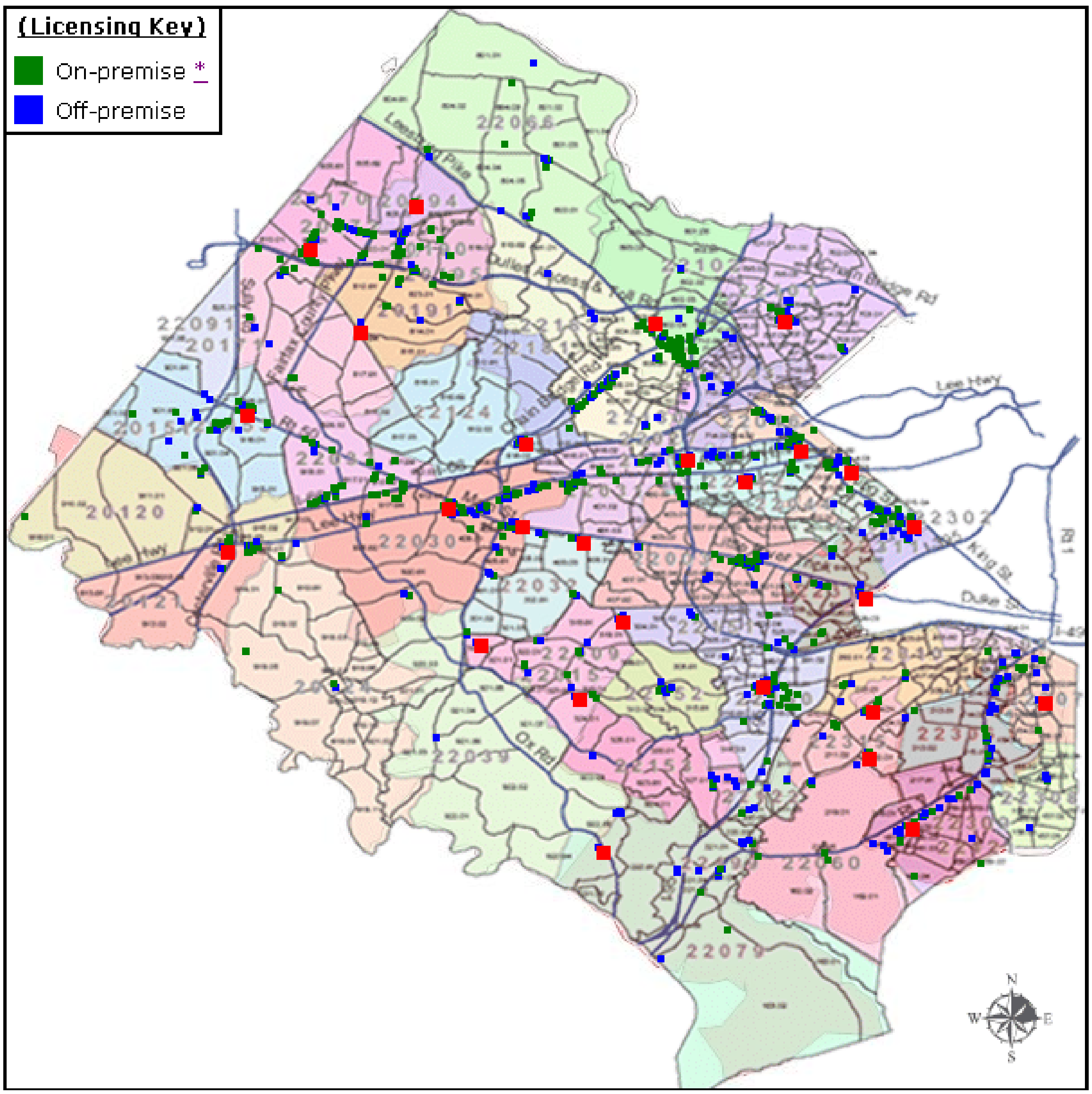


Figure 17: Fairfax County map with two overlays, street map, ABC stores and alcohol sales establishments

It is very apparent from this image, for instance, that there is a large cluster of alcohol establishments along the length of the Route 1 corridor.

The user may wish to retrieve additional demographic information about an individual zip code region. To do this, simply select the desired zip code region, for example:

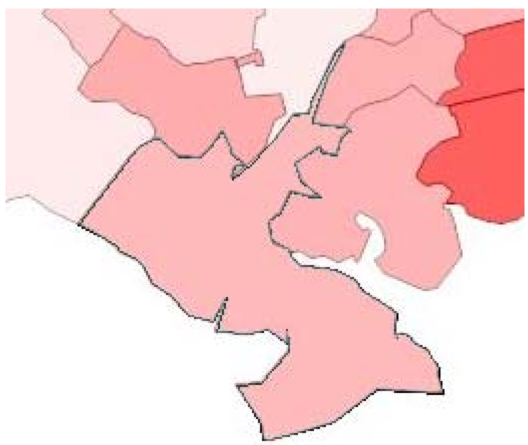


Figure 18: The Lorton zip code map.

And demographic and zip code specific information is instantly displayed in a side panel. The proportion of each race and the number of alcohol establishments are displayed for the particular zip code. These may be adjusted within the zip code to see zip code specific effects rather than countywide effects. The adjustment is done by “widgets”, the << or >> buttons on either side of the displayed data. The race (both in absolute population increase and balance of ethnic distribution) and the number of alcohol establishments may be adjusted up or down. By clicking on the Update but the simulation for the whole county with the adjusted data may be examined. Addition demographic data may be examined by scrolling down. Data on ethnicity (race), age distribution, household size, family size, and gender by ethic groups are available.

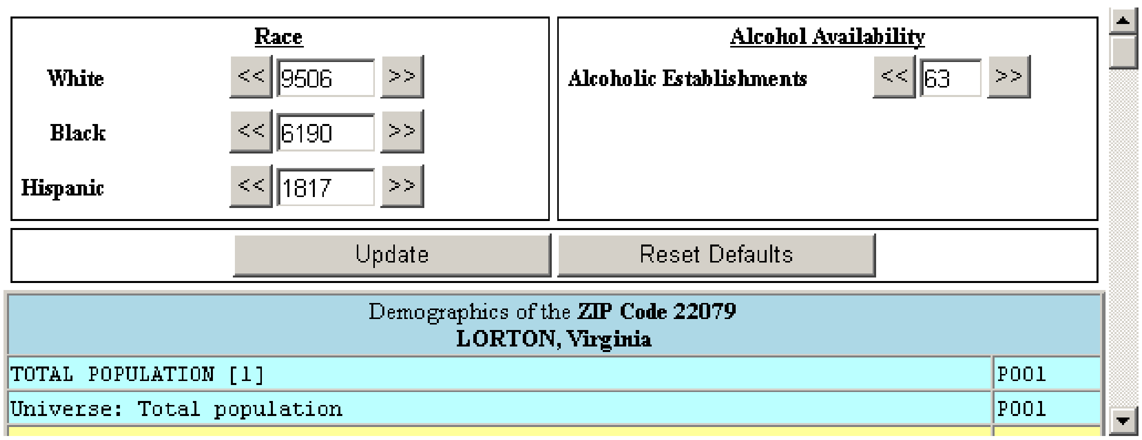


Figure 19: A partial view of the Lorton demographic statistics window.

Yet, another portion of the interface allows one to modify the probabilities countywide used during the simulation. For example if an ethnicity adjustment is made in this mode, the individual zip code data are suppressed and the population proportions are taken county wide. However, the probabilities associated with the alcohol availability within zip codes are not modified. Low availability is defined as those zip codes having fewer than 16 alcohol establishments. High availability includes those zip codes having 50 or more alcohol establishments. Moderate availability includes all the rest. As in the case with ethnicity (race), if an availability adjustment is made in this mode, the individual zip code data are suppressed and the alcohol availability is taken countywide. However, the probabilities associated with ethnicity within a zip code are not modified. The Fairfax County map is then redrawn accordingly to reflect its new outcome status. This can be done using these “widgets”:

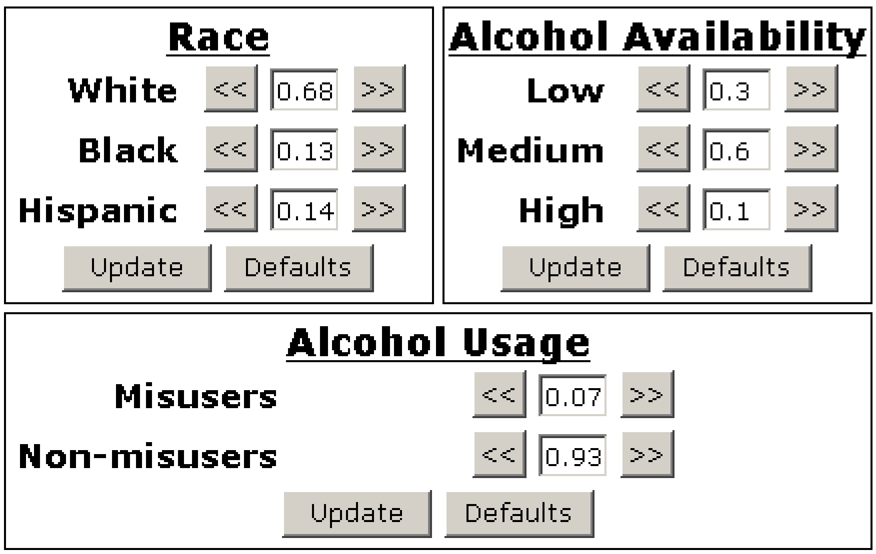
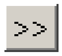


Figure 20: A screen shot of the controller window.

“Misusers” are defined as individuals who are either alcohol abusers or alcohol dependent as defined in the NLAES data. By default, the values displayed are the probabilities used by the simulation based on actual data. The user can modify these values by pressing on the << and >> buttons. For example, if the



button for the White race is pressed 2 times, the following probabilities are displayed:

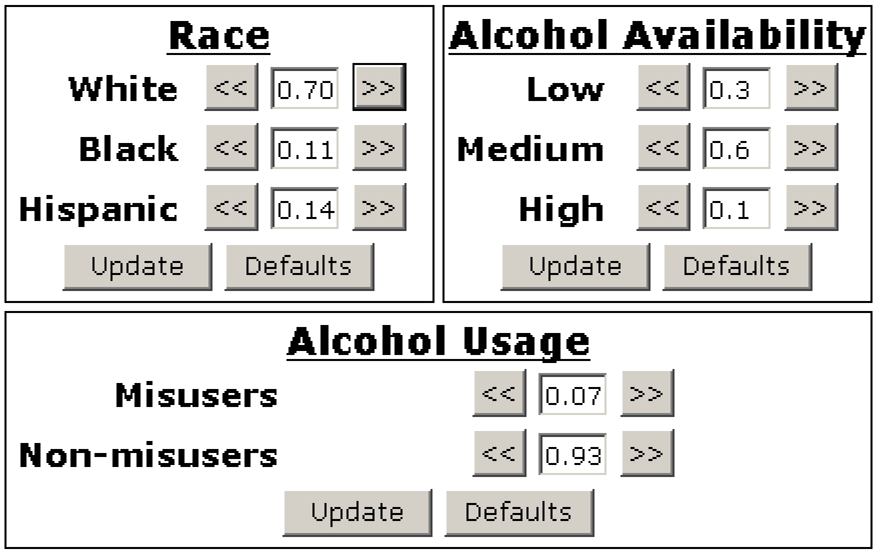


Figure 21: Adjusted probabilities in the controller window.

To “commit” this update, and re-run the simulation, press the


button within the race box.

The simulation is regenerated after a few seconds, and the page updated to the initial page displaying the Fairfax County color-coded map with population pie charts overlaid:

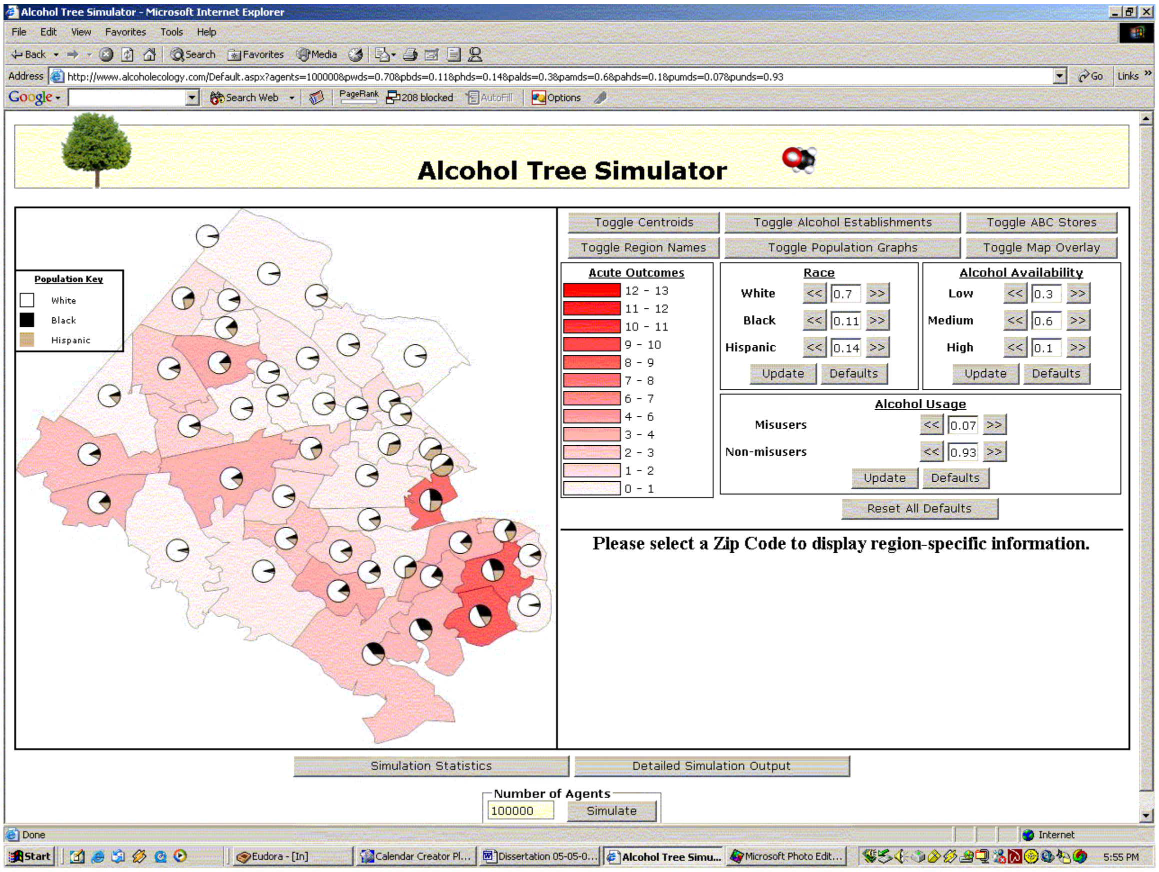


Figure 22: Screen shot of the updated application after altered probabilities.

The same method can be used to change the probability values for Alcohol Availability and Alcohol Usage. To return to the default values for one particular probability group (i.e. Race), use the Defaults button. To reload all default probability values at once, one can use the: Reset All Defaults button.

As an example, when the Low alcoholic availability probability is set to 1.0 (meaning Medium and High availability values are set to 0), our rendered Fairfax map changes from:

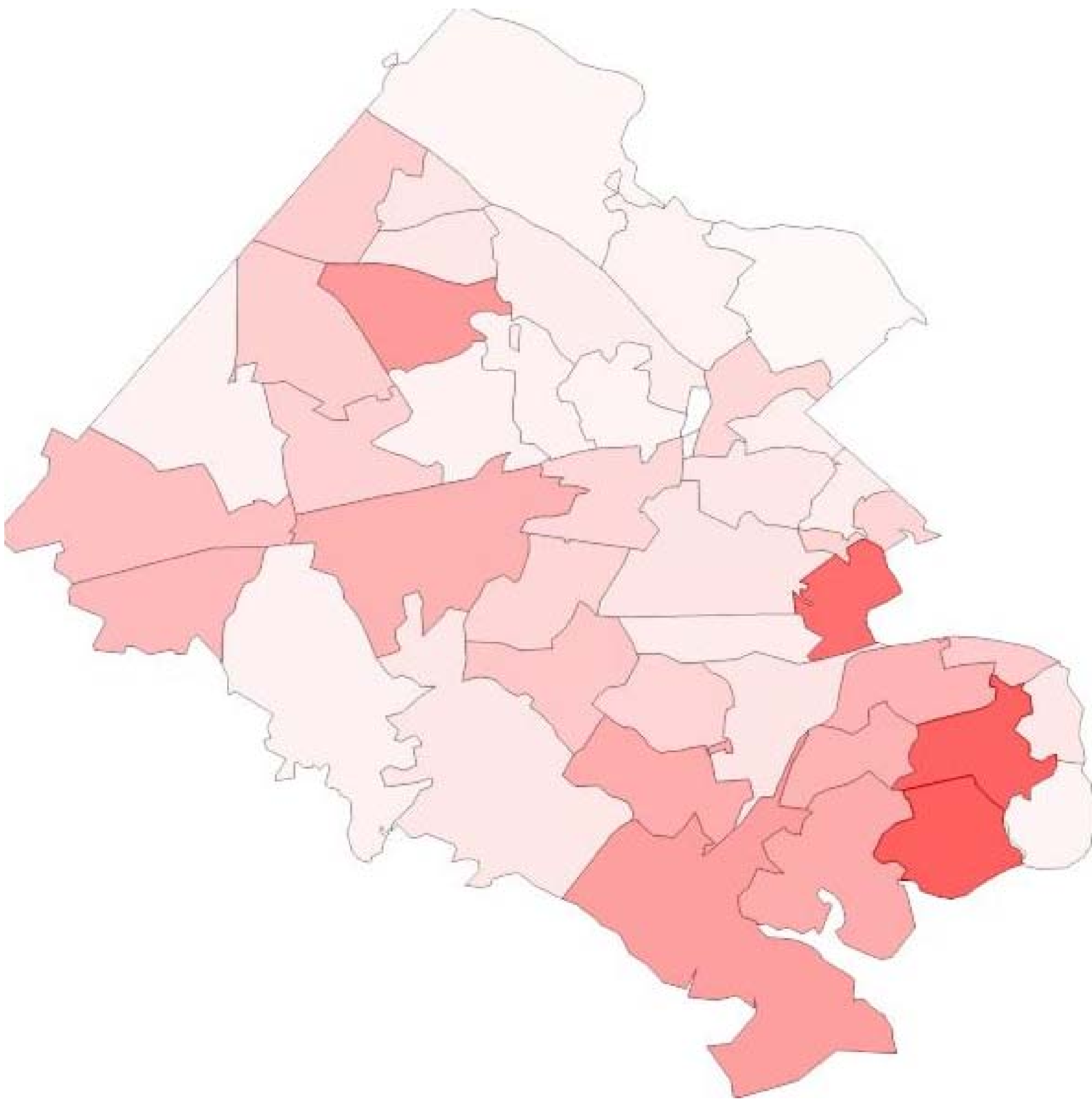


Figure 23: Fairfax map showing intensity of acute outcomes with default probabilities based on actual data.

to (with changed low alcohol availability value):

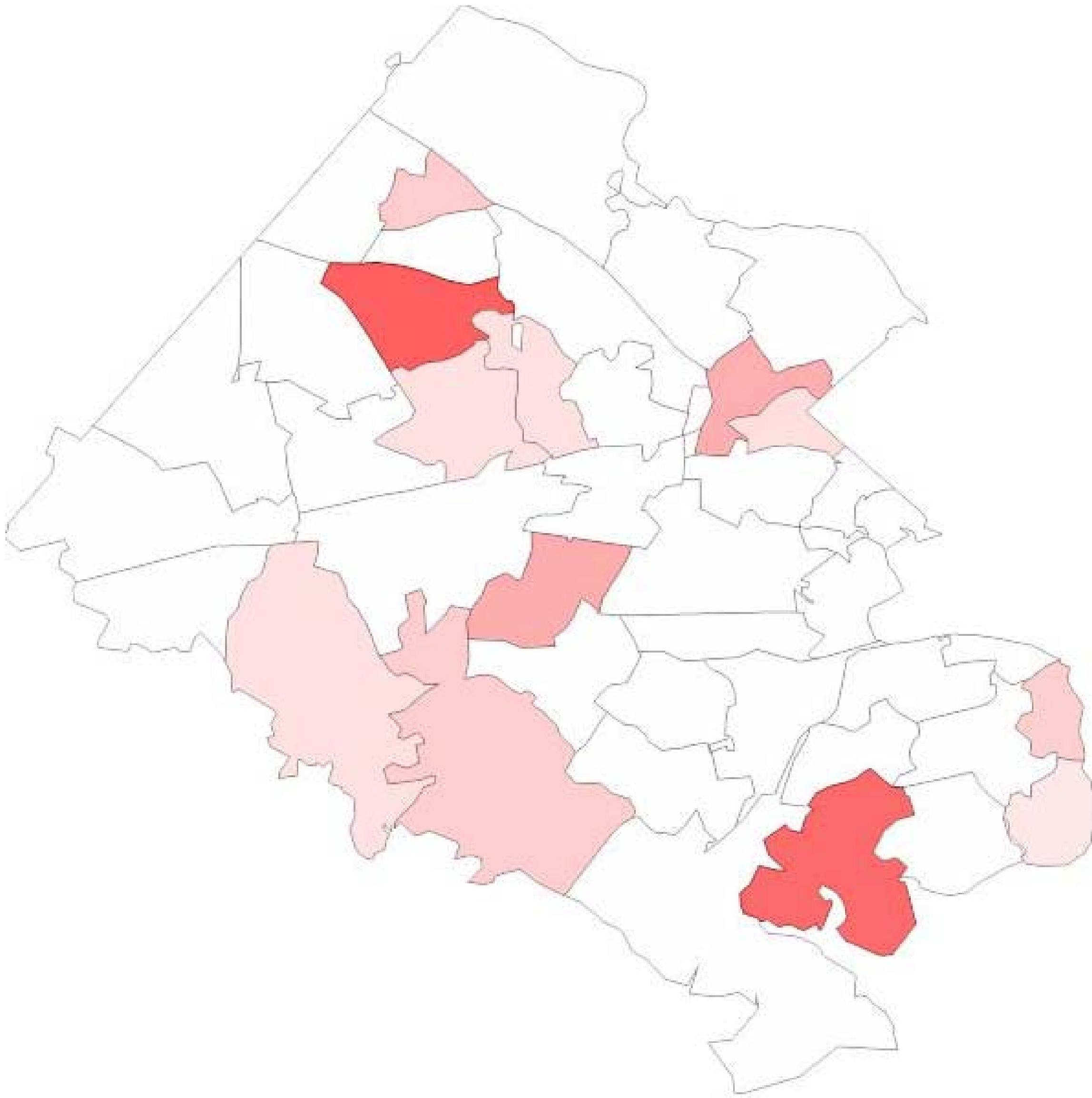
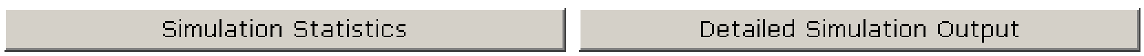


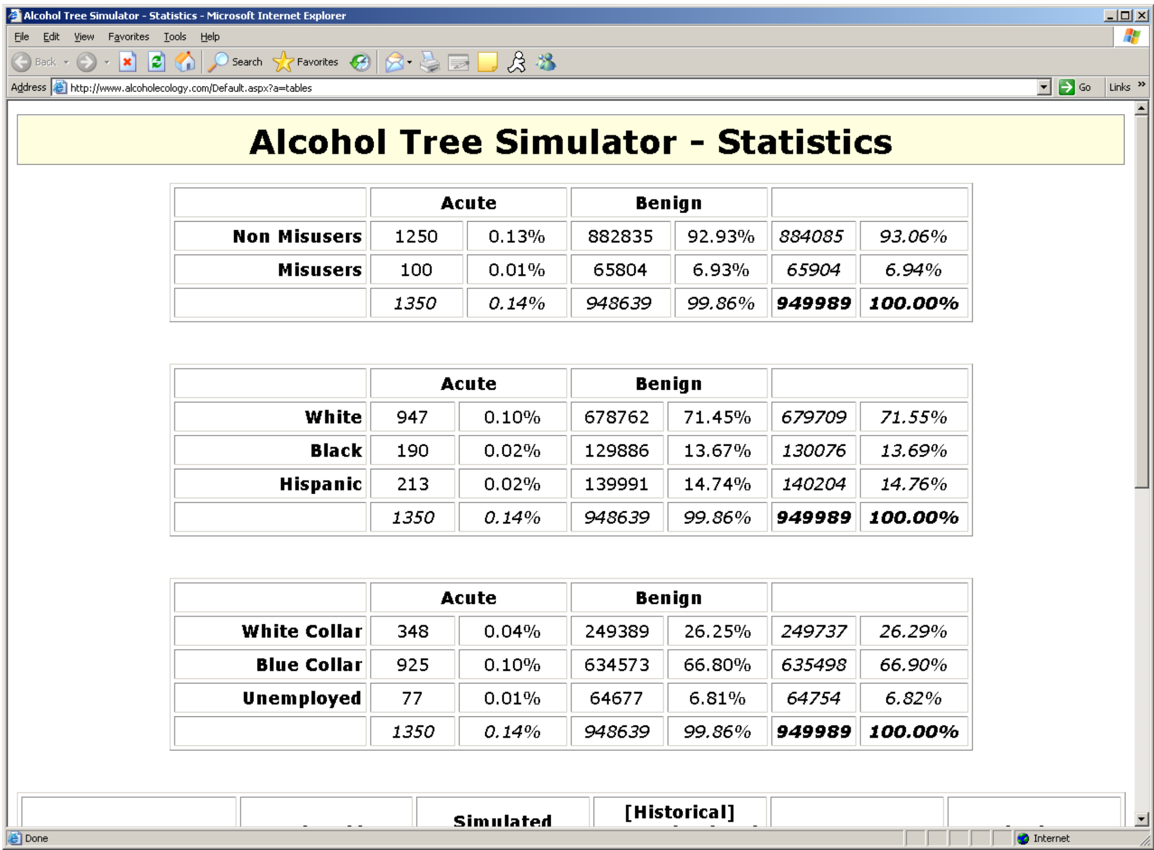
Figure 24: Fairfax County map with only low outlet availability in each zip code.

In this image, the darkest red regions correspond to Fort Belvoir and Reston. A dark red region indicates one with a high number of acute outcomes, whereas a white region indicates a region with none.

To view additional statistics and detailed simulation output, the buttons located at the bottom of the page are used:



Simulation Statistics are popped up in a new window, as follows:



Alcohol Tree Simulator - Statistics

	Acute		Benign			
Non Misusers	1250	0.13%	882835	92.93%	*884085*	*93.06%*
Misusers	100	0.01%	65804	6.93%	*65904*	*6.94%*
	1350	*0.14%*	*948639*	*99.86%*	***949989***	***100.00%***

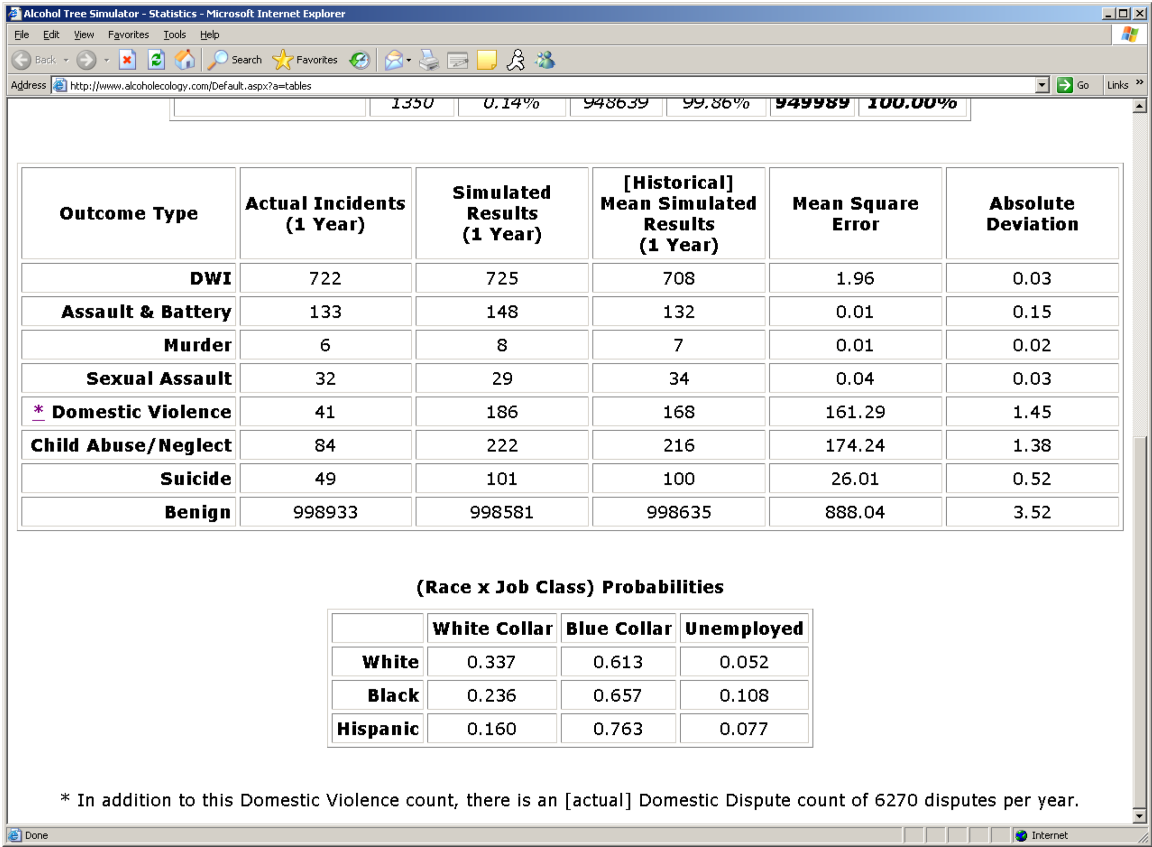
	Acute		Benign			
White	947	0.10%	678762	71.45%	*679709*	*71.55%*
Black	190	0.02%	129886	13.67%	*130076*	*13.69%*
Hispanic	213	0.02%	139991	14.74%	*140204*	*14.76%*
	1350	*0.14%*	*948639*	*99.86%*	***949989***	***100.00%***

	Acute		Benign			
White Collar	348	0.04%	249389	26.25%	*249737*	*26.29%*
Blue Collar	925	0.10%	634573	66.80%	*635498*	*66.90%*
Unemployed	77	0.01%	64677	6.81%	*64754*	*6.82%*
	1350	*0.14%*	*948639*	*99.86%*	***949989***	***100.00%***

		Simulated	[Historical]		

Figure 25: Simulation statistics window.

Acute outcomes are any of the outcomes DWI, assault and battery, murder, suicide, sexual assault, child abuse/neglect, or domestic violence. A benign outcome occurs if none of the acute outcomes happen. More statistics are available by scrolling to the bottom of the page:



	1350	0.14%	948639	99.86%	**949989**	**100.00%**

Outcome Type	Actual Incidents (1 Year)	Simulated Results (1 Year)	[Historical] Mean Simulated Results (1 Year)	Mean Square Error	Absolute Deviation
DWI	722	725	708	1.96	0.03
Assault & Battery	133	148	132	0.01	0.15
Murder	6	8	7	0.01	0.02
Sexual Assault	32	29	34	0.04	0.03
* **Domestic Violence**	41	186	168	161.29	1.45
Child Abuse/Neglect	84	222	216	174.24	1.38
Suicide	49	101	100	26.01	0.52
Benign	998933	998581	998635	888.04	3.52

(Race x Job Class) Probabilities

	White Collar	Blue Collar	Unemployed
White	0.337	0.613	0.052
Black	0.236	0.657	0.108
Hispanic	0.160	0.763	0.077

* In addition to this Domestic Violence count, there is an [actual] Domestic Dispute count of 6270 disputes per year.

Figure 26: Additional simulation statistical tables.

The first three two way tables display acute and benign outcomes based on various criteria, for instance, alcohol usage, ethnicity, and job class. The fourth table displays each outcome type and its associated number of actual incidents per year (from a police data file), its simulated results (from the current simulation), a historical mean simulated result (an average taken from 100 previously run simulations), the historical mean square error, and the historical average absolute deviation. The Table on this page shows the ethnicity (Race) by Job Class probabilities. These probabilities were obtained from the Bureau of Labor Statistics. The Bureau of Labor Statistics had a relatively fine set of job classes. I aggregated these into white collar, blue collar and unemployed. These job class mappings were then used to generate our final table, discussed earlier, and shown below:

	White Collar	Blue Collar	Unemployed
White	0.337	0.613	0.052
Black	0.236	0.657	0.108
Hispanic	0.160	0.763	0.077

Figure 27: Screen shot of joint probabilities.

This last table contains probability values. This window can now be closed, and the initial Fairfax County map interface is again displayed. The Detailed Simulation Output button may be used to popup another new window, this time displaying – as the title states – detailed simulation output. The output generated is multiple pages long, showing the entire alcohol “tree” as generated by the simulation. The tree hierarchy can be seen, along with the branching of acute and benign outcomes for different job classes, alcohol availabilities, and alcohol usage classes. This output is shown in Table 16 in Appendix B.

Once this window is closed, one may again return to the main Fairfax County map page. If one wishes to re-run the simulation at any time, one may enter the number of agents to simulate for (or leave the default value set) and again press the simulate button, shown below:

Chapter 7: Conclusions, Contributions and Future Work

7.1 Conclusions

This dissertation presents a methodology and a tool for investigating the impact of interventions on the distribution of acute outcomes associated with alcohol misuse. This work only begins to suggest a strategy. However, I believe that a longer-term effort building on this methodology has the potential for becoming a valuable public policy tool. In my work, I suggest the possibility of investigating the effects of adjusting appropriate conditional probabilities to discover the effect of the adjustment on the multinomial distribution of acute and benign outcomes. The second part of the policy tool would be to create interventions that achieve the appropriate adjustments in the conditional probabilities. There is still much work to be done.

The grand vision for this work involves the complex interaction among various sectors of society reminiscent of a biological ecosystem. Some have argued that ecosystem and ecology are misnomers and words like environment should be used. However, I believe that the analogy with ecosystem is highly appropriate in that competition for resources and prey-predator analogies are easily imagined. Indeed, ecology has been used in the literature of the National Institute of Alcohol Abuse and Alcoholism. See for example Gruenewald et al. (2002) and Jacob and Johnson (1997).

7.2 Contributions

To briefly outline my contributions in this dissertation, I constructed a model framework involving a stochastic, time dependent, spatially oriented stochastic digraph model. I undertook a substantial data collection effort from a wide variety of sources. Based on limitations of that data collection effort, I constructed and implemented a reduced model, which did not model time-of-day effects, but captured some geo-spatial effects. Data for time-of-day effects are not currently available because of the need to have an extensive data collection effort on activities on both weekdays and weekends in order to have a realistic scenario development portfolio. Surrogate data are also not readily available. Based on data I did collect, I undertook estimation of transition probabilities in my digraph model. I implemented the model in JAVA with simulation and "what if" capabilities and created a website to showcase my implementation. I provided some statistical assessments and did an approximate calibration of the model through the simulation.

Having articulated these component contributions, I believe they aggregate in to a systematic approach to exploring and ultimately suppressing acute outcomes. This approach can be developed further into a policy tool that has the potential for positively impacting the misuse of alcohol and the acute consequences associated with that misuse. As outlined in the abstract alcohol misuse is a major problem leading to both societal and personal acute and chronic outcomes.

7.3 Future Work

This dissertation work was intended to be a proof-of-concept and should be considered a work in progress. I want to extend this work in several directions. I plan to incorporate time of day, weekend vs. weekday, and gender considerations. I plan to expand to include Arlington County (and perhaps all of Northern Virginia) and give special consideration to underage and college age drinkers. I presently do not have any age effects incorporated in the model, but it is well known that underage and college age drinking is a societal problem and that the typical peak drinking period in life is during late adolescence and early adulthood. The presence of a large university and a teenage and young adult population in Northern Virginia provides an excellent opportunity to investigate impact on these populations. I plan to include improved scenario development through activity generation studies. I plan to include additional ethnic and job categories. I plan to include mobility information. I plan to include a more complex social network (including remediation facilities and treatment facilities) and try to build longer-term adaptation to interventions. I plan to improve the ability to explore interventions/adjustments within local units (zip codes) and indeed to explore finer scale and more homogeneous geo-spatial structures.

References

References

Alcohol Problems in Later Life. (1998). A Pacific Northwest Extension Publication. Report PNW32, <http://eesc.orst.edu/agcomwebfile/edmat/pnw342.pdf>.

The American Heritage Dictionary of the English Language. (2004). Fourth Edition, Boston: Houghton Mifflin Company.

The American Heritage Stedman's Medical Dictionary. (2002). Boston: Houghton Mifflin Company.

Asylum Seeker Resource Centre. (2005). Asylum Seeker and Temporary Protection Visa Rights Guide. <http://asrc.ewock.com/images/Textfile326.doc#_Toc53197043>.

Australian Institute of Criminology. Australian Government. (2005). Illicit Drugs and Alcohol. <http://www.aic.gov.au/research/drugs/types/alcohol.html>.

Answers.com. (2005). <http://www.answers.com/topic/blood-alcohol-content>.

Butler Center for Research Hazelden Foundation. (2000). “Alcohol use and domestic violence” Hazelden Foundation. November Issue.

Canadian Centre for Occupational Health and Safety. (1999). “OSH Answers – What does LD50 mean?” <www.ccohs.ca/oshanswers/chemicals/**ld50**.html>.

Changeux, J. and Edelman, G. (2001). “Drug use and abuse,” The Brain. Somerset, N.J.: Transaction Publishers.

Clinical and Experimental Research. (2005). Journal on Alcoholism. March Issue

Dawson D.A. (2000). “Alternative Measures and Models of Hazardous Consumption,” Journal of Substance Abuse. 12(1-2): 79- 91.

De Freitas, J. F. G., Niranjan, M., and Gee, A. H. (2000). “Hierarchical Bayesian models for regularization in sequential learning.” Neural Computation. 12(4): 933-953.

Dunham, Will (2004) “Military has a growing drinking problem,” Reuters. <http://news.myway.com/top/article/id/389850%7Ctop%7C03-08-2004::16:49%7Creuters.html>

Fillmore M.T., Carscadden J.L., Vogel-Sprott M. (1998) “Alcohol, cognitive impairment and expectancies,” Journal of Studies on Alcohol. 59(2):174-179.

Fogarty, Jennifer N. (2002). “Cognitive processes and motor skills differ in sensitivity to alcohol impairment,” Journal of Studies on Alcohol. 63(4): 404-411.

Gerhard Gmel, and Rehm, Jürgen. (2003). “Harmful alcohol use,” NIAAA. <http://www.niaaa.nih.gov/publications/arh27-1/52-62.htm>

Gilmore, Ian. (2004). “Alcohol advertising - consultation document,” Royal College of Physicians. London, issued 23 September.

Gruenewald, P. J., Remer, L. and Lipton, R. (2002). “Evaluating the alcohol environment: Community geography and alcohol problems,” <http://www.niaaa.nih.gov/publications/arh26-1/42-48.htm>

Harrison, Ric. (1997). “Alcohol impaired driving statistics,” Community Alcohol Information Program. <www.nh-dwi.com/caip-206.htm>.

Huntingdonshire District Council. (2005). “Drugs And alcohol,” <www.huntsdc.gov.uk/Community+and+People/Crime+and+disorder/Crime/Drugs+and+Alcohol.htm>.

Iowa Substance Abuse Information Center. (2005). “Violence prevention,” <http://www.drugfreeinfo.org/tdomestic.html>.

Jabob, T. and Johnson, S. (1997). “Parenting influences on the development of alcohol abuse and dependence,” Alcohol Health and Research World, 21(3), 204-209.

Leonard, K. E. and Blane, H. T. (eds.) (1999). Psychological Theories of Drinking and Alcoholism, 2nd Edition, New York: Guilford Press.

Moss, H.B., and Tarter, R.E. (1993). “Substance abuse, aggression, and violence,” American Journal on Addiction. 2(2):149-160.

Nahas, Gabriel G., Burks, T. F. (1997). Drug Abuse in the Decade of the Brain. Amsterdam: IOS Press.

NIAAA. (2000) 10th Special Report to Congress on Alcohol and Health, NIAAA, <http://www.niaaa.nih.gov/publications/10report/intro.pdf>

NIAAA. (2005). "Alcohol, violence, and aggression," Alcohol Alert. <http://alcoholism.about.com/cs/alerts/l/blnaa38.htm>.

Petrucci, R. H., Harwood, W. S., and Herring, F. G. (2001). General Chemistry: Principles and Modern Applications. 8th Edition. Upper Saddle River, N.J.: Prentice Hall.

Pharmacology Central. (2000). <http://dwb.unl.edu/Teacher/NSF/C10/C10Links/www.pharmcentral.com/sedatives.htm>.

Reiss, A.J., Jr., and Roth, J.A. (1994). Understanding and Preventing Violence. Vol. 3. Washington, DC: National Academy Press.

Rutgers University. (2004). A Guide to Alcohol. <http://health.rutgers.edu/brochures/alcohol.htm>.

Shakhashiri, B. Z. (2005). "The chemical of the week," <http://scifun.chem.wisc.edu/CHEMWEEK/ETHANOL/ethanol.html>.

Snijders, Tom. (January 2001). "The statistical evaluation of social network dynamics," in Sociological Methodology 2001. (Sobel, M. E. and Becker, M.P., eds.). London: Basil Blackwell.

University of Ottawa, Health Services. (2005). <http://www.uottawa.ca/health/information/addictions-alcohol.html>.

U.S. Department of Health and Human Services and SAMHSA's National Clearinghouse for Alcohol and Drug Information. (1995)."Domestic violence and alcohol and other drugs," <http://www.health.org/govpubs/ml001/>.

The Virginia Alcohol Safety Action Program (2005), <http:www.vasap.state.va.us/>.

The Virginia Department of Alcoholic Beverage Control. (2005). <http://www.abc.state.va.us>.

Wikipedia. (2005). <http://en.wikipedia.org/wiki/Prohibition>.

Wilson, R., and Dufour, M., (eds.) (2000) Epidemiology of Alcohol Problems in Small Geographic Areas. NIAAA Research Monograph No. 36. NIH Pub. No. 00–4357. Bethesda, MD: National Institute on Alcohol Abuse and Alcoholism.

World Health Organization (2002). The World Health Report 2002 - Reducing Risks, Promoting Healthy Lives. Geneva: World Health Organization.

Zweben, Joan E, and Denning, Pat. (1998). The Alcohol and Drug Wildcard: Substance Use and Psychiatric Problems in People with HIV. UCSF AIDS Health Project Monograph Series. No. 2. San Francisco: AHP.

Appendix A: Hierarchical Statistical Estimation of Probabilities

The model is built around Fairfax County in Northern Virginia. One key element is to divide the population into relatively homogeneous sub-populations based on the idea that their behaviors with respect to alcohol will be relatively homogeneous. Even so, it is desirable to model variability into such a sub-population.

Let D be the digraph representing the possible states of an individual, for example, traveling to work or in a bar or at home, working or engaged in recreation, sober or inebriated. As discussed previously, each simulated agent x moves randomly among the vertices of D according to agent-dependent, time-varying transition probabilities

$$P_{mn}(x;t) = \text{Prob}\{\text{Next node is } n \mid \text{Current node is } m \text{ (and all other past information)}\} \quad \text{(A.1)}$$

Ideally I would like to create a full model of the behavior of every individual. However, such a modeling effort entails estimation of far too many parameters relative to available data.

At the same time, however, variability among agents is essential. As a compromise, in future work I intend to employ Bayesian hierarchical models. Specifically, I define a manageably small number of classes of individuals representing different socio-demographic and geographical characteristics. (One example would be single college students living in Western Fairfax County.) Individuals in class X have randomly varying, but statistically identical transition matrices, so that one needs only to estimate a prior distribution on transition matrices associated with each class.

Ignoring time dependence for the moment, the probabilities $P_{mn}(x)$ are selected randomly from a distribution over transition matrices. For both modeling flexibility and technical reasons, the model employs class-dependent Dirichlet prior distributions for each row of the transition matrix. That is, with $x \in \mathrm{X}$ and m fixed, the probabilities $P(\cdot) = P_{m\cdot}(x)$ will be sampled from a Dirichlet distribution

$$f\left(p, \beta^{\mathrm{X},m}\right) = \frac{1}{\Psi\left(\beta^{\mathrm{X},m}\right)} \prod_{n} P_{n}^{\;u_{n}^{\mathrm{X},m} - 1} \tag{A.2}$$

where $\Psi\left(\beta^{\mathrm{X},m}\right) = \prod_{n} \Phi\left(\beta_{n}^{\mathrm{X},m}\right) / \Phi\left(\sum_{n} \beta_{n}^{\mathrm{X},m}\right)$ and $\beta^{\mathrm{X},m}$ are class-dependent parameters to be estimated. Mathematically, the Dirichlet distribution is the conjugate prior for the multinomial probabilities represented by $P_{m\cdot}(x)$.

An extreme alternative to this hierarchical approach is to assume that all individuals in a class have exactly the same transition probabilities: $P_{m\cdot}(x) \equiv P_{m\cdot}(\mathrm{X})$ for all $x \in \mathrm{X}$. This alternative does not reduce the number of parameters to be estimated, but does reduce significantly the richness of the simulation. Time dependence can be handled similarly. One can quantize time, for example into {Morning, Noon-time, Early Afternoon, Late afternoon, Evening, Night}, which leads to Dirichlet hyperparameters $\beta^{\mathrm{X},m,\tau}$, where τ is a time interval. In effect, this inserts time into the modeling hierarchy.

In principle, it is also necessary to model the sojourn times that an agent spends in various vertices of the digraph. Markov renewal processes, which allow such distributions to depend (only) on the current and next states, are a potential tool. However, estimation of class- and time-dependent (hyper- or not) sojourn distributions is clearly impossible from currently available data. Minimally, I impose a Markov assumption that sojourn times are exponentially distributed and depend only on the current state (which would allow use of hierarchical models for sojourn times as well as states), but it is virtually certain that even stronger assumptions would be needed.

A.1 Estimation

The hierarchical model for transition probabilities contained in (A.1) and (A.2) requires estimation of

$$\mathrm{K}_{\mathrm{X}} \times |L|^2 \times \mathrm{K}_{\tau}$$

Dirichlet hyperparameters, where K_{X} is the number of classes, $|L|^2$ is the number of vertices in the digraph and K_{τ} is the number of intervals into which time is quantized. Plausible values for these are $\mathrm{K}_{\mathrm{X}} = 10, L = 20$ and $\mathrm{K}_{\tau} = 6$, leading to 24,000 parameters!

Although estimation of individual Dirichlet hyperparameters $\beta^{\mathrm{X},m,\tau}$ is straightforward (relevant count data may be used directly), the estimation of this many parameters from diverse and limited data is a significant challenge. Strategies are available, however, which include, for example:

• Additional assumptions, for example, that for some vertices $m, \beta^{\mathrm{X},m,\tau}$, does not depend on X or τ.

• Structural assumptions on the digraph, which force some of the $P_{mn}(\mathrm{X},\tau)$ to be zero, (i.e. some or even many transitions may be impossible).

• Use of expert opinion to provide values of $\beta^{\mathrm{X},m,\tau}$ for which data are nonexistent or too weak.

An extreme, but potentially viable assumption for sojourn time distributions would be that the distribution of the sojourn time in vertex m is an exponential distribution depending only on m, and not on the agent's class or the time of day. In this case, only as many exponential parameters (or associated Gamma distribution hyperparameters) would need to be estimated. The most significant problem to remain, however, is characterization and quantification of the uncertainties associated with various estimated values and propagation of these to uncertainties in the output of the simulation.

A.2 Simulation

In part to characterize uncertainties, and because closed form computation of probabilities of adverse outcomes is not feasible, it is necessary to perform (multiple) simulations of the entire population of Fairfax County. The principal steps in this process are:

i. For each individual agent x, determine its class $\mathrm{X}(x)$. This may be done either deterministically or stochastically.

ii. For each agent in class X and each time interval τ, stochastically generate transition probabilities $P_{mn}(\mathrm{X},\tau)$ using the (A.2) Dirichlet distributions with hyper-parameters $\beta^{\mathrm{X}(x),\cdot,\tau}$. (A more efficient implementation would generate $\beta^{\mathrm{X}(x),m,\tau}$ only if x ever enters vertex m during time interval τ; it is not clear at this stage whether such details matter.)

iii. Similarly, stochastically generate sojourn time distributions $D_{mn}(x,\tau)$.

iv. Using the stochastically generated transition probabilities $P_{mn}(x,\tau)$ and sojourn time distributions $D_{mn}(x,\tau)$ simulate and record the day-long path of agent x through the digraph.

The computational effort is, to first approximation, linear in the number of agents and quadratic in the number of vertices in the digraph. This is feasible. The effort is also linear, obviously, in the number of replications of the simulation. No improvement on this is possible.

A.3 Some Possible Simplifications

One strategy for modeling transition probabilities is to augment (A.2) with dependence on covariates. This kind of modeling can be useful in modeling the dependence of transitions on covariates such as age, gender, time of day, etc. Several kinds of discrete choice models are possible. For simplicity, consider the multinomial logistic model for k choices specified as follows. For agent x, suppose there are covariate vectors $v_{xm1},\cdots,v_{xmk}$ and a parameter vector w such that the probability of making choice n, $1 \le n \le k$, given state m is

$$P_{mn}(x) = \frac{e^{\varphi'_{xmn}w}}{\sum_{i=1}^{k} e^{\varphi'_{xmi}w}}$$

As a simple example, consider the following model. Suppose an agent has finished work and must make a choice between three activities, (1) go home, (2) stop for a drink, or (3) do some other leisure activity such as eating or shopping away from home. The probability model depends on two other factors, the agent's gender and time of day. For simplicity, time of day is discretized to six periods. Then a simple model could be coded as

$$\varphi'_{xmi}w = x_{1i}w_1 + x_{2i}w_2 + x_{3i}w_3 + l_x w_4 + d_{x1}w_5 + d_{x2}w_6 + d_{x3}w_7$$

where the x_{ni} terms are dummy variables for activities,

$$x_{ni} = 1, n = i \text{ and } x_{ni} = 0, n \neq i\,.$$

Similarly, $l_x = 1$ if agent x is male and zero otherwise is a dummy variable for gender, and the t_{xn} are dummy variables for three of the four time-of-day periods.

More generally, this discrete choice model can be used as part of a hyperprior specification for the Dirichlet prior of (A.2). Under (A.2), the prior mean for choice n is

$$E\left(P_{mn}(x) \mid \beta^{X,m}\right) = \frac{e^{\beta_n^{X,m}}}{\sum_i e^{\beta_i^{X,m}}}.$$

Therefore, the parameters of the Dirichlet prior for agent x could be modeled as

$$\beta_n^{X,m} = \varphi'_{xmn} w + \varepsilon_{xn}$$

for suitable independent mean zero random variables ε_{xn} and agent x in class X. I believe judicious use of this model may simplify the number of transition probabilities to be estimated.

Estimates of a portion of these probabilities can be obtained from trip and activity data sets. For example, the choice model without the drinking choice may be estimated from such a data set. One property of the simple multinomial logistic choice model is independence of irrelevant alternatives. To be specific, consider the simplified choice model

$$P_{mn}(x) = \frac{e^{\beta_n}}{\sum_{i=1}^{k} e^{\beta_i}}.$$

For any subset $N \subset \{1, \ldots, k\}$, it is easy to show that the conditional probability of choice n given $n \in N$ is simply

$$P_{mn}(x \mid n \in N) = \frac{e^{\beta_n}}{\sum_{i \in N} e^{\beta_i}}.$$

This means that certain parameters can be estimated from marginal data. One could estimate some parameters in these models from the Portland data set without drinking behavior, and add parameters for drinking alternatives to match conditional probabilities from other surveys.

Appendix B: Data Tables

Table 4: Zip codes and percentages within Fairfax County.

Zip code	Place	County	% in FFC	Zip code Population	Fairfax Population
20120	CENTREVILLE	FX and Loudoun	99.70%	34,372	34,269
20121	CENTREVILLE	FX	100.00%	26,043	26,043
20124	CLIFTON	FX	100.00%	14,863	14,863
20151	CHANTILLY	FX	100.00%	17,162	17,162
20170	HERNDON	FX	100.00%	38,114	38,114
20171	HERNDON	FX	100.00%	33,811	33,811
20190	RESTON	FX	100.00%	14,258	14,258
20191	RESTON	FX	100.00%	28,881	28,881
20194	RESTON	FX	100.00%	12,666	12,666
22003	ANNANDALE	FX	100.00%	53,029	53,029
22015	BURKE	FX	100.00%	42,938	42,938
22027	DUNN LORING	FX	100.00%	1,377	1,377
22030	FAIRFAX	FX and City of FX	55.70%	39,431	21,963.00
22031	FAIRFAX	FX and City of FX	95.80%	26,768	25,644
22032	FAIRFAX	FX and City of FX	95.70%	28,886	27,644
22033	FAIRFAX	FX	100.00%	33,043	33,043
22039	FAIRFAX STATION	FX	100.00%	18,471	18,471
22041	FALLS CHURCH	FX	100.00%	26,966	26,966
22042	FALLS CHURCH	FX and City of Falls Church	98.00%	30,505	29,895
22043	FALLS CHURCH	FX	100.00%	22,815	22,815
22044	FALLS CHURCH	FX and City of Falls Church	93.20%	13,109	12,218
22046	FALLS CHURCH	FX and City of Falls Church	31.10%	14,252	4,432.00
22060	FORT BELVOIR	FX	100.00%	8,100	8,100
22066	GREAT FALLS	FX and Loudoun	96.90%	16,958	16,432
22079	LORTON	FX	100.00%	19,607	19,607
22101	MC LEAN	FX and Arlington	99.60%	28,433	28,319
22102	MC LEAN	FX	100.00%	18,607	18,607
22124	OAKTON	FX	100.00%	13,867	13,867
22150	SPRINGFIELD	FX	100.00%	22,208	22,208
22151	SPRINGFIELD	FX	100.00%	16,587	16,587

22152	SPRINGFIELD	FX	100.00%	28,236	28,236
22153	SPRINGFIELD	FX	100.00%	33,177	33,177
22180	VIENNA	FX	100.00%	22,021	22,021
22181	VIENNA	FX	100.00%	14,527	14,527
22182	VIENNA	FX	100.00%	22,983	22,983
22302	ALEXANDRIA	FX and City of Alexandria	6.50%	15,014	976
22303	ALEXANDRIA	FX	100.00%	14,268	14,268
22304	ALEXANDRIA	FX and City of Alexandria	0.70%	40,763	285
22306	ALEXANDRIA	FX	100.00%	28,249	28,249
22307	ALEXANDRIA	FX	100.00%	9,505	9,505
22308	ALEXANDRIA	FX	100.00%	12,554	12,554
22309	ALEXANDRIA	FX	100.00%	29,259	29,259
22310	ALEXANDRIA	FX	100.00%	26,082	26,082
22311	ALEXANDRIA	FX and City of Alexandria	7.00%	16,665	1167
22312	ALEXANDRIA	FX and City of Alexandria	76.50%	28,370	21,703
22315	ALEXANDRIA	FX	100.00%	21,901	21,901

Table 5: Zip code population and demographic information.

Zip code	Hispanic or Latino	White alone	Black or African American alone
20120	2,287	24,793	2,371
20121	2,648	15,860	2,619
20124	511	12,200	409
20151	1,785	11,424	893
20170	7,117	21,669	3,130
20171	1,892	24,226	1,751
20190	1,724	9,287	1,408
20191	3,467	18,459	2,939
20194	544	10,163	681
22003	7,175	31,154	2,923
22015	3,593	29,220	2,061
22027	43	1,000	42
22030	3,812	26,290	2,877
22031	3,698	14,991	1,415
22032	1,593	21,410	1,151
22033	1,802	23,667	1,635
22039	530	15,103	673
22041	9,104	10,152	2,568
22042	6,795	15,710	1,410
22043	3,182	14,203	1,099
22044	3,760	6,106	658
22046	1,543	10,853	440
22060	786	4,322	2,419
22066	437	14,220	324
22079	1,817	9,506	6,190
22101	1,113	22,988	473
22102	1,013	13,355	508
22124	559	11,654	398
22150	4,143	10,919	1,597
22151	1,596	10,998	713
22152	2,323	19,363	1,622
22153	2,312	22,583	3,155
22180	1,995	15,892	745
22181	946	11,230	397
22182	810	17,549	659
22207	2,205	24,010	1,418
22302	1,593	10,781	1,409
22303	2,416	8,438	1,982
22304	6,099	19,311	9,771
22306	5,471	11,928	7,541
22307	712	7,686	650
22308	324	11,385	312

22309	4,231	13,987	8,161
22310	2,563	17,091	2,829
22311	2,607	7,326	4,213
22312	6,101	10,918	6,145
22315	1,442	14,756	2,866

Table 6: ABC Store gallons sold, and gross sale figures, in dollars.

20121	14151 Germain Drive	53,196	$3,449,672	$569,164	$2,880,507
20151	13944 Lee Jackson Highway	53,526	$3,341,019	$550,548	$2,790,471
20170	1238 Elden St (Herndon)	45,475	$2,941,960	$486,208	$2,455,753
20171	13300 Franklin Farm Road (Herndon)	3,897	$245,049	$40,140	$204,909
20171	2555 John Milton Drive (Herndon)	30,674	$1,906,781	$314,535	$1,592,246
20194	1454 N Point Village SC (Reston)	44,227	$2,954,114	$486,987	$2,467,127
22003	7200 Little River Tnpk., East	40,411	$2,466,360	$407,894	$2,058,467
22015	9574 Old Keene Mill Road	29,292	$1,773,680	$292,215	$1,481,465
22015	5739 Burke Centre Parkway	20,247	$1,213,777	$200,226	$1,013,552
22030	10308 Willard Way	1,091	$65,523	$10,744	$54,778
22041	3556-E S. Jefferson St.	40,087	$2,525,831	$415,610	$2,110,221
22042	7263 Arlington Blvd	1,256	$83,318	$13,618	$69,700
22042	8105 Lee Highway	32,237	$2,023,522	$332,735	$1,690,787
22044	6198 Arlington Boulevard	26,126	$1,591,564	$262,087	$1,329,476
22101	1446 Chain Bridge Road	49,673	$3,277,889	$540,167	$2,737,722
22124	2928 Chain Bridge Road	1,006	$64,262	$10,484	$53,778
22150	6400 Springfield Plaza	43,707	$2,699,522	$446,268	$2,253,254
22151	8966 Burke Lake Road	27,913	$1,588,288	$262,226	$1,326,063
22180	436 East Maple Ave (Vienna)	47,601	$2,884,970	$473,365	$2,411,605
22182	8520 Tyco Road (Vienna) (conv)	93,230	$7,612,546	$1,263,521	$6,349,025
22307	1524 Belle View Boulevard	35,377	$2,073,786	$340,799	$1,732,987
22309	8628 Richmond Highway	49,969	$3,226,636	$534,318	$2,692,318
22309	7778 Gunston Plaza	16,580	$988,012	$163,153	$824,859
22315	5926 Kingstowne Center	39,633	$2,541,710	$418,273	$2,123,436

Table 7: Alcohol Establishment license information, and status.

Zip Code	County	Status	Privileges	Trade Name
22180	FAIRFAX COUNTY	Active	Annual Caterers License	Marco Polo Caterers
22150	FAIRFAX COUNTY	Active	Annual Caterers License	R & R Catering
22043	FAIRFAX COUNTY	Active	Annual Caterers License	Doubletree Hotel
22043	FAIRFAX COUNTY	Active	Annual Caterers License	Stratford College
22042	FAIRFAX COUNTY	Active	Annual Caterers License	Fairview Park Marriott
22031	FAIRFAX COUNTY	Active	Annual Caterers License	RSVP Catering
22030	FAIRFAX CITY	Active	Annual Caterers License	@ George Mason University
20190	FAIRFAX COUNTY	Active	Annual Caterers License	Hyatt Regency Reston
22102	FAIRFAX COUNTY	Active	Annual Caterers License	Compass Group USA Inc
22101	FAIRFAX COUNTY	Active	Annual Caterers License	Sodexho Management Inc
22182	FAIRFAX COUNTY	Active	Annual Mixed Beverage Special Event	The Barns
22182	FAIRFAX COUNTY	Active	Annual Mixed Beverage Special Event	Wolf Trap Foundation
22121	FAIRFAX COUNTY	Active	Annual Mixed Beverage Special Event	Woodlawn Plantation
22151	FAIRFAX COUNTY	Active	Beer Imported – In-State	Guiffre Distributing Company
22150	FAIRFAX COUNTY	Active	Beer Imported – In-State	Premium Distributors of Virginia
22079	FAIRFAX COUNTY	Active	Beer Imported – In-State	Balearic Beverage Distributors
22079	FAIRFAX COUNTY	Active	Beer Imported – In-State	Hop And Wine Beverages LLC
22079	FAIRFAX COUNTY	Active	Beer Imported – In-State	Service Distributing
22032	FAIRFAX COUNTY	Active	Beer Imported – In-State	Stronbol
20191	FAIRFAX COUNTY	Active	Beer Imported – In-State	American Fidelity Trading Ltd
20190	FAIRFAX COUNTY	Active	Beer Imported – In-State	Martlet Importing Co
20151	FAIRFAX COUNTY	Active	Beer Imported – In-State	Premium Distributors Of Virginia
20151	FAIRFAX COUNTY	Active	Beer Imported – In-State	Select Wines

20151	FAIRFAX COUNTY	Active	Beer Imported – In-State	King Wholesale Inc
22079	FAIRFAX COUNTY	Active	Beer Imported – In-State	Dionysos Imports Inc
22079	FAIRFAX COUNTY	Active	Beer Imported – In-State	Eastern Wholesale
22079	FAIRFAX COUNTY	Active	Beer Imported – In-State	Ithaka Imports Inc
22079	FAIRFAX COUNTY	Active	Beer Imported – In-State	Young Won Trading Inc
22031	FAIRFAX COUNTY	Active	Beer Imported – In-State	Suprex International Ltd
20190	FAIRFAX COUNTY	Active	Beer Imported – In-State	Fosters USA LLC
22309	FAIRFAX COUNTY	Active	Beer Off Premises	Rorers Produce
22180	FAIRFAX COUNTY	Active	Beer Off Premises	CVS/Pharmacy 1847
22180	FAIRFAX COUNTY	Active	Beer Off Premises	Vienna Shell
22180	FAIRFAX COUNTY	Active	Beer Off Premises	Vienna Tiger Mart
22152	FAIRFAX COUNTY	Active	Beer Off Premises	7 Eleven Store 2581 30479
22150	FAIRFAX COUNTY	Active	Beer Off Premises	El Mercado
22150	FAIRFAX COUNTY	Active	Beer Off Premises	Interfuel LLC
22046	FALLS CHURCH CITY	Active	Beer Off Premises	Go Cong Supermarket
22046	FAIRFAX COUNTY	Active	Beer Off Premises	Hillwood Mart
22041	FAIRFAX COUNTY	Active	Beer Off Premises	Duangrat Oriental Food Mart
22041	FAIRFAX COUNTY	Active	Beer Off Premises	San Miguel Market
22039	FAIRFAX COUNTY	Active	Beer Off Premises	Exxon #25741
22033	FAIRFAX COUNTY	Active	Beer Off Premises	Circle K Mobil 5885
22031	FAIRFAX CITY	Active	Beer Off Premises	Phuoc Loc Market
22031	FAIRFAX CITY	Active	Beer Off Premises	Citgo 7 Eleven Store 2581 30474
22030	FAIRFAX COUNTY	Active	Beer Off Premises	Popes Head Market
22030	FAIRFAX CITY	Active	Beer Off Premises	Fairfax Texaco Shell
22030	FAIRFAX CITY	Active	Beer Off Premises	E & C VA 010

Table 8: Alcohol availability outlets, color-coded by outlet type.

Zip Code	Quantity	ABC Quantity	City
22027	1		Dunn Loring
22181	1		Vienna
22304	3		Alexandria
20194	4	1	Reston
22060	5		Fort Belvoir
20124	6		Clifton
22308	6		Alexandria
22039	8		Fairfax Station/Fairfax
22307	8	1	Alexandria
22124	10	1	Oakton
22046	11		Falls Church
22043	12		Falls Church
22032	14		Fairfax
20191	15		Reston
22152	16		Springfield/West Springfield
22153	16		Springfield
20120	19		Centreville
22066	21		Great Falls/Fairfax
22151	21	1	Springfield
22303	22		Alexandria
22312	25		Alexandria
22315	25	1	Alexandria/Fairfax

20121	27	1	Centreville
22031	28		Fairfax
22310	28		Alexandria/Franconia
22015	29	2	Burke
22182	34	1	Vienna/Tysons Corner/Fairfax
20171	35	2	Herndon
22306	35		Alexandria/Groveton
22044	36	1	Falls Church/Fairfax
22309	40	2	Alexandria
22030	45	1	Fairfax
22033	46		Fairfax/Chantilly
22102	48		McLean/Tysons Corner

20190	50		Reston/Herndon/Fairfax
20151	51	1	Chantilly
22101	52	1	McLean/Fairfax
22041	62	1	Falls Church/Baileys Crossroads/Fairfax/Lorton
22079	63	1	Lorton/Fairfax/Mason Neck
20170	65	1	Herndon
22180	66	1	Vienna
22042	68	2	Falls Church/Fairfax
22003	69	1	Annandale
22150	78	1	Springfield/Fairfax

Table 9: Leading causes of death

PLANNING DISTRICT AND CITY / COUNTY	DEATHS FROM ALL CAUSES		UNINTENTIONAL INJURY		SUICIDE		CHRONIC LIVER DISEASE		HOMICIDE AND LEGAL INTERVENTION	
	NUMBER	RATE	NUMBER	RATE	NUMBER	RATE	NUMBER	RATE	NUMBER	RATE
COMMONWEALTH OF VIRGINIA	56,095	886.9	2,357	34.8	770	10.8	571	8.3	449	6.2
PLANNING DISTRICT 1	1,142	1,079.0	59	62.7	28	29.4	12	11.4	7	7.9
LEE COUNTY	301	1,088.7	20	84.1	6	25.8	4	14.7	2	8.4
SCOTT COUNTY	317	1,020.3	16	62.5	12	43.4	1	4.7	3	14.1
WISE COUNTY	455	1,067.9	20	50.3	10	25.1	4	8.8	2	4.8
NORTON CITY	69	1,486.4	3	69.8			3	65.0		
PLANNING DISTRICT 2	1,229	979.6	65	53.9	25	19.6	18	13.5	11	9.4
BUCHANAN COUNTY	264	1,050.7	16	53.8	3	11.5	6	22.0	2	8.2
DICKENSON COUNTY	166	916.2	10	60.9	3	18.1	7	36.1	1	6.3
RUSSELL COUNTY	310	1,003.0	18	60.1	8	24.5	2	6.0	5	16.8
TAZEWELL COUNTY	489	951.4	21	46.3	11	22.6	3	5.8	3	7.0
PLANNING DISTRICT 3	2,192	935.3	92	44.9	31	15.1	19	8.4	7	3.3
BLAND COUNTY	80	1,070.0	4	61.9						
CARROLL COUNTY	293	808.8	11	30.1	7	24.4	2	6.2	1	3.5
GRAYSON COUNTY	223	1,003.5	19	99.2	2	9.6	1	3.9		
SMYTH COUNTY	394	971.6	12	34.4	5	14.8	3	7.9	1	2.8
WASHINGTON COUNTY	453	772.1	25	47.1	11	18.6	4	5.8	4	6.4
WYTHE COUNTY	312	931.9	11	38.2	2	5.6	4	11.8		
BRISTOL CITY	300	1,176.5	8	41.4	4	22.2	3	18.4		
GALAX CITY	137	1,396.4	2	24.6			2	27.3	1	17.5
PLANNING DISTRICT 4	1,364	917.7	68	44.1	17	10.7	13	8.7	6	3.0
FLOYD COUNTY	132	751.7	13	91.6			2	11.6		
GILES COUNTY	202	990.3	11	68.8	6	32.6	2	9.9		
MONTGOMERY COUNTY	482	830.3	20	31.2	3	2.5	7	11.7	2	2.8
PULASKI COUNTY	426	1,047.8	21	57.4	4	10.3	2	4.8	2	6.4
RADFORD CITY	122	1,085.2	3	28.8	4	40.9			2	3.8
PLANNING DISTRICT 5	3,038	934.5	99	34.0	49	17.4	36	11.7	16	6.0
ALLEGHANY COUNTY	123	806.5	3	21.9	1	5.3	2	12.6		
BOTETOURT COUNTY	246	861.6	14	51.4	5	15.8	2	5.1	2	5.1
CRAIG COUNTY	40	769.6					1	17.4		
ROANOKE COUNTY	928	882.7	31	33.7	17	18.2	12	11.7	1	0.9
CLIFTON FORGE CITY	91	1,119.9	4	55.8						
COVINGTON CITY	112	1,162.8	3	41.2	2	32.1	2	26.2		
ROANOKE CITY	1,168	985.5	31	28.0	21	20.9	13	12.8	12	13.4
SALEM CITY	330	1,066.6	13	46.7	3	9.2	4	14.4	1	4.4
PLANNING DISTRICT 6	2,372	851.6	86	31.6	28	10.8	21	7.6	8	3.1
AUGUSTA COUNTY	468	719.9	13	20.8	5	6.9	5	6.5	1	1.6
BATH COUNTY	74	1,220.0	6	117.5	1	17.9				
HIGHLAND COUNTY	27	723.5	2	53.8						
ROCKBRIDGE COUNTY	178	790.7	15	72.3	2	11.1	2	7.9		
ROCKINGHAM COUNTY	539	739.8	21	31.2	5	7.1			2	3.0
BUENA VISTA CITY	88	1,102.1	4	52.5	1	19.7				
HARRISONBURG CITY	296	948.1	12	35.9	2	7.0	4	15.2	2	8.2
LEXINGTON CITY	123	1,410.7	1	8.7	2	39.2	1	10.7		
STAUNTON CITY	313	953.3	4	15.0	7	28.4	4	14.1		

WAYNESBORO CITY	266	1,049.8	8	35.2	3	15.3	5	22.3	3	14.9
PLANNING DISTRICT 7	1,660	868.5	67	36.1	26	13.7	23	11.3	6	3.2
CLARKE COUNTY	130	908.2	1	11.3	1	10.0	2	13.6		
FREDERICK COUNTY	373	752.7	18	32.0	5	9.0	8	13.8	2	3.7
PAGE COUNTY	233	871.2	12	47.0	1	3.3	2	7.5	1	4.5
SHENANDOAH COUNTY	363	818.6	17	48.8	8	22.5	4	8.9	1	3.2
WARREN COUNTY	301	999.9	11	37.4	6	19.7	4	12.0	1	3.3
WINCHESTER CITY	260	1,021.1	8	32.5	5	21.2	3	11.7	1	4.6
PLANNING DISTRICT 8	8,448	726.2	359	24.4	139	8.0	77	5.3	50	2.8
ARLINGTON COUNTY	1,102	698.0	34	17.3	15	8.2	5	2.8	7	3.7
FAIRFAX COUNTY	4,210	673.1	186	24.1	69	7.4	34	4.2	26	2.7
LOUDOUN COUNTY	695	827.5	34	30.7	10	9.1	5	4.7	1	0.9
PRINCE WILLIAM COUNTY	1,130	888.6	57	29.5	32	11.2	12	6.5	10	3.3
ALEXANDRIA CITY	833	792.6	26	21.6	6	4.9	11	10.1	5	4.3
FAIRFAX CITY	140	640.5	3	15.0	2	9.6	3	13.5		
FALLS CHURCH CITY	71	612.1	2	18.9	2	20.8	2	18.4		
MANASSAS CITY	215	1,262.0	14	51.1	2	4.7	2	12.2	1	7.5
MANASSAS PARK CITY	52	1,511.6	3	75.6	1	21.2	3	51.1		

Note: Age adjusted rates are derived by applying age group population relatives from the standard 2000 U.S. Census of population to the current age-specific rates of the same population groups. They are more comparable between populations with different age distributions than are un-adjusted rates based on their populations. The National Center for Health Statistics suggests the use of the 2000 population relatives for use by the states for comparative purposes. The adjusted rates lie below the unadjusted rates after the base year as a population ages.

Table 10: Resident alcohol induced deaths by race and sex
by underlying causes of death
Virginia, 2000

ALCOHOL INDUCED DEATHS BY THE NATIONAL CENTER FOR HEALTH STATISTICS CLASSIFICATION WITH ICD 10 CODES	TOTAL DEATHS	WHITE RACE MALE	WHITE RACE FEMALE	BLACK RACE MALE	BLACK RACE FEMALE	ALL OTHER RACES MALE	ALL OTHER RACES FEMALE
TOTAL ALCOHOL INDUCED DEATHS	352	208	59	68	14	3	
Alcohol Acute Intoxication (F100)	6	2	1	3			
Alcohol Harmful Use .... (F101)	30	16	5	8	1		
Alcohol Dependence Syndrome (F102)	61	41	6	11	3		
Alcohol Other Behavioral Disorders (F103-F109)	15	7	2	4	1	1	
Alcoholic Degeneration of Nervous System (G312)	4	1	1	2			
Alcoholic Polyneuropathy (G621)	1	1					
Alcoholic Cardiomyopathy (I426)	7	4	1	2			
Alcoholic Gastritis (K292)	2	2					
Alcoholic Liver Disease (K70)	208	121	40	36	9	2	
Excessive Blood Level of Alcohol (R780)							
Accidental Poisoning by Alcohol (X45)	18	13	3	2			
Intentional Self-Poisoning by Alcohol (X65)							
Undetermined Intent Poisoning by Alcohol (Y15)							

NOTE: The Table Excludes Deaths Due To Accidents, Homicides, And Other Causes Indirectly Related To Alcohol.

Table 11: Resident Alcohol Induced Deaths for Fairfax County
By Zip Code and Race/Sex
Virginia, 2000

Zip Code	Total Deaths	White Race		Black Race		Other Race	
		Male	Female	Male	Female	Male	Female
20120	1	1	0	0	0	0	0
20190	1	0	1	0	0	0	0
22003	1	0	1	0	0	0	0
22015	3	0	0	1	0	1	1
22021	2	0	0	0	0	1	1
22041	4	1	0	1	0	1	1
22046	1	1	0	0	0	0	0
22101	1	1	0	0	0	0	0
22303	1	1	0	0	0	0	0
22306	1	1	0	0	0	0	0
22309	1	1	0	0	0	0	0
22310	2	2	0	0	0	0	0
22312	1		1	0	0	0	0
All	**20**	**9**	**3**	**2**	**0**	**3**	**3**

Note: The table excludes deaths due to accidents, homicides, and other causes indirectly related to alcohol.
All zip codes not included in this table had no alcohol-related deaths.

Table 12: A sample of motor vehicle crashes

VEH-DOCUMENT-NUMBER	VEH-NUMBER	VEH-LICENSED-STATE	VEH-INSURANCE-INDICATOR	VEH-TYPE	VEH-AGE	VEH-ESTIMATED-SPEED	VEH-MANUVER	VEH-POINT-OF-IMPACT	VEH-DAMAGE	VEH-REPAIR-COST	VEH-DL-CDL
3391097	2	34	1	5	2	50	1	5	8	100	2
542066	1	47	1	2	7	20	1	1	6	4000	0
542066	2	47	1	2	1	35	1	3	1	5000	0
542067	1	47	1	2	14	999	1	1	1	2500	1
542067	2	47	1	2	24	999	8	5	1	1000	1
462172	1	47	1	1	10	3	7	8	8	1000	1
462172	2	47	1	1	6	25	1	3	8	1500	1
462176	1	47	1	1	15	10	3	1	6	4000	1
462176	2	47	1	1	99	20	1	2	8	3000	1
542527	1	47	1	1	2	0	1	4	1	4000	1
542076	1	47	3	3	10	35	1	1	1	1500	1
542076	2	47	1	1	5	5	1	5	1	500	1
542080	1	47	1	1	5	15	4	1	8	2000	1
542080	2	47	1	2	1	25	1	1	8	1500	1
530233	1	47	1	5	4	0	15	1	6	500000	2
331763	1	47	1	1	1	45	1	1	8	6000	1
331763	2	47	1	1	10	10	3	6	8	2000	1
332720	1	33	1	1	13	20	3	1	8	800	1
332720	2	47	1	4	7	35	1	6	8	200	1
380217	1	47	1	3	2	20	1	8	8	6000	1
380217	2	34	1	1	10	25	3	8	8	4000	1
560386	1	47	1	2	4	55	1	2	8	4000	1
462680	1	52	3	0	99	40	9	5	1	100	0
120020	1	47	1	1	1	80	10	1	6	19000	1
120021	1	47	1	1	6	45	1	1	8	2000	1
120025	1	52	2	3	12	45	1	1	8	2000	1
321357	1	52	1	1	15	0	8	5	8	300	1

Table 13: Sample of crime statistics

Case No	Date	Subcensus	Patrol Area	Final Event Type
2000139002643	5/18/00	40202	440	Aggravated Assault
2000254000280	9/10/00	15102	210	Aggravated Assault
2000276000097	10/2/00	52504	431	Aggravated Assault
2000157001679	6/5/00	50901	400	Aggravated Assault
2000260000248	9/16/00	52103	600	Aggravated Assault
2000057000322	2/26/00	15401	211	Aggravated Assault
2000092001988	4/1/00	21501	220	Aggravated Assault
2000325001262	11/20/00	21101	631	Aggravated Assault
2000203002742	7/21/00	16101	240	Aggravated Assault
2000248001983	9/4/00	82201	510	Aggravated Assault
2000186001811	7/4/00	81901	511	Aggravated Assault
2000169001996	6/17/00	82101	511	Aggravated Assault
2000336002262	12/1/00	82301	530	Aggravated Assault
2000332002205	11/27/00	82301	530	Aggravated Assault
2000284002090	10/10/00	81104	520	Aggravated Assault
2000234002488	8/21/00	15401	211	Aggravated Assault
2000186002165	7/4/00	15401	211	Aggravated Assault
2000317000438	11/12/00	51501	410	Aggravated Assault
2000058000200	2/27/00	51501	410	Aggravated Assault
2000155000148	6/3/00	40202	440	Aggravated Assault
2000328001380	11/23/00	51603	411	Aggravated Assault
2000212000162	7/30/00	15102	210	Drinking while driving
2000165000009	6/13/00	15403	211	Drinking while driving
2000171000394	6/19/00	90104	110	Drinking while driving
2000321000057	11/16/00	22401	630	Drug Event
2000140000065	5/19/00	81001	521	Drug Event
2000309001962	11/4/00	15201	210	Drug Event
2000214000025	8/1/00	60504	330	Drug Event
2000208002076	7/26/00	15501	230	Drug Event
2000033001402	2/2/00	21001	620	Drug Event
2000353001728	12/18/00	80502	510	Drug Event
2000237001107	8/24/00	70602	320	Drug Event
2000204000981	7/22/00	81001	521	Drug Event
2000059000119	2/28/00	15201	210	Drug Event
2000064000134	3/4/00	32201	711	Drug Event
2000042001823	2/11/00	51501	410	Drug Event
2000095002063	4/4/00	15301	210	Drug Event
2000232001714	8/19/00	91304	101	Drug Event
2000259000031	9/15/00	61202	830	Drug Event
2000296001918	10/22/00	61501	831	Drunk in Public
2000334000509	11/29/00	61501	831	Drunk in Public
2000002000600	1/2/00	50101	341	Drunk in Public

2000201000060	7/19/00	15401	211	Drunk in Public
2000249002162	9/5/00	15401	211	Drunk in Public
2000329000019	11/24/00	15401	211	Drunk in Public
2000152001273	5/31/00	61601	831	Drunk in Public
2000152002382	5/31/00	61701	331	Drunk in Public
2000330000183	11/25/00	61701	331	Drunk in Public
2000347002153	12/12/00	40201	331	Drunk in Public
2000124001524	5/3/00	50501	400	Drunk in Public
2000143001961	5/22/00	50501	400	Drunk in Public
2000154002017	6/2/00	51501	410	Drunk in Public
2000279002296	10/5/00	51501	410	Drunk in Public
2000316002198	11/11/00	51501	410	Drunk in Public
2000020001587	1/20/00	51501	410	Drunk in Public
2000023000472	1/23/00	51501	410	Drunk in Public
2000048001823	2/17/00	51501	410	Drunk in Public
2000091000713	3/31/00	51501	410	Drunk in Public
2000100001953	4/9/00	51501	410	Drunk in Public
2000107000009	4/16/00	51401	401	Drunk in Public
2000113001960	4/22/00	51501	410	Drunk in Public
2000188002172	7/6/00	51501	410	Drunk in Public
2000259002493	9/15/00	51501	410	Drunk in Public
2000276002339	10/2/00	51501	410	Drunk in Public
2000293000008	10/19/00	51401	401	Drunk in Public
2000294002151	10/20/00	51401	401	Drunk in Public
2000334002280	11/29/00	51501	410	Drunk in Public
2000340001589	12/5/00	51501	410	Drunk in Public
2000163000284	6/11/00	50204	342	Drunk in Public
2000030000185	1/30/00	51501	410	Drunk in Public
2000358001895	12/23/00	22101	640	DWI 2nd
2000360000032	12/25/00	51503	420	DWI 2nd
2000360000124	12/25/00	70403	311	DWI 2nd
2000192002337	7/10/00	61102	831	DWI 3rd
2000228001741	8/15/00	22301	630	DWI 3rd
2000184000042	7/2/00	32801	730	DWI 3rd
2000110000060	4/19/00	21004	620	DWI 3rd
2000221001391	8/8/00	15102	210	DWI 3rd
2000243000019	8/30/00	21003	620	DWI 3rd
2000260000039	9/16/00	40504	820	DWI 3rd
2000263000046	9/19/00	71302	340	DWI 3rd
2000217002555	8/4/00	20102	620	DWI 3rd w/ < 5
2000217002349	8/4/00	60602	330	DWI 3rd w/in 5-10 yrs
2000295002111	10/21/00	50702	441	DWI 3rd w/in 5-10 yrs
2000192000889	7/10/00	50501	400	Embezzlement
2000088001257	3/28/00	51504	420	Embezzlement
2000340002000	12/5/00	31801	710	Embezzlement

2000321000679	11/16/00	50304	342	Embezzlement
2000217000746	8/4/00	50702	441	Embezzlement
2000024001442	1/24/00	80205	302	Embezzlement
2000121001490	4/30/00	80205	302	Embezzlement
2000322001851	11/17/00	80205	302	Embezzlement
2000340000837	12/5/00	61604	340	Embezzlement
2000096001172	4/5/00	61605	331	Embezzlement
2000343000648	12/8/00	50103	341	Embezzlement
2000279001245	10/5/00	61701	331	Embezzlement
2000337001714	12/2/00	40201	331	Illegal Poss Alcohol <21
2000339000735	12/4/00	61802	830	Illegal Poss Alcohol <21
2000343002533	12/8/00	82203	511	Illegal Poss Alcohol <21
2000344002033	12/9/00	22101	640	Illegal Poss Alcohol <21
2000345000088	12/10/00	92001	121	Illegal Poss Alcohol <21
2000347000205	12/12/00	80303	310	Illegal Poss Alcohol <21
2000351000044	12/16/00	82301	530	Illegal Poss Alcohol <21
2000351002011	12/16/00	22201	740	Illegal Poss Alcohol <21
2000362002035	12/27/00	91304	101	Illegal Poss Alcohol <21
2000365001837	12/30/00	52105	601	Illegal Poss Alcohol <21
2000273002454	9/29/00	80503	510	Liquor - Illegal Purch/Poss Minor
2000330001833	11/25/00	70802	320	Liquor - Illegal Purch/Poss Minor
2000106000203	4/15/00	81902	511	Liquor - Illegal Purch/Poss Minor
2000010000068	1/10/00	81102	521	Liquor - Illegal Purch/Poss Minor
2000348000176	12/13/00	15403	211	Poss open alcohol container
2000177001391	6/25/00	52801	421	Possess alcohol in park
2000209002152	7/27/00	30801	721	Possess alcohol in park
2000049000621	2/18/00	80502	510	Purchase/Poss Tobacco
2000102001256	4/11/00	80502	510	Purchase/Poss Tobacco
2000125001014	5/4/00	80502	510	Purchase/Poss Tobacco
2000270001612	9/26/00	80502	510	Purchase/Poss Tobacco
2000280000544	10/6/00	82201	510	Purchase/Poss Tobacco
2000290001420	10/16/00	80502	510	Purchase/Poss Tobacco
2000315001847	11/10/00	80502	510	Purchase/Poss Tobacco
2000063000847	3/3/00	80901	940	Purchase/Poss Tobacco
2000343000850	12/8/00	91702	820	Purchase/Poss Tobacco
2000156000566	6/4/00	81901	511	Purchase/Poss Tobacco
2000344000084	12/9/00	82101	511	Purchase/Poss Tobacco
2000171000581	6/19/00	70802	320	Purchase/Poss Tobacco
2000291001107	10/17/00	70405	311	Purchase/Poss Tobacco
2000215001376	8/2/00	52001	430	Robbery
2000175002267	6/23/00	52301	700	Robbery
2000091000997	3/31/00	30402	721	Robbery
2000074000930	3/14/00	20401	200	Robbery
2000087001525	3/27/00	21002	620	Robbery
2000132001513	5/11/00	21401	201	Robbery

2000132002418	5/11/00	51501	410	Robbery
2000095000714	4/4/00	52601	601	Robbery
2000077002401	3/17/00	21001	620	Robbery
2000331001352	11/26/00	21001	620	Robbery
2000333002200	11/28/00	31601	611	Robbery
2000364000073	12/29/00	52503	431	Robbery
2000161001849	6/9/00	31401	731	Robbery
2000364001193	12/29/00	52104	600	Robbery

Table 14: Conditional probability of being an alcohol misuser given ethnicity, job class and zip code in Fairfax County

Zip Code	Ethnicity	White Collar	Blue Collar	Unemployed
22066	White	0.033489369	0.127971358	0.008904672
	African American	0.000534097	0.003186973	0.0003432
	Hispanic	0.000657953	0.004132662	0.000356032
22102	White	0.029618776	0.113180842	0.007875499
	African American	0.000748201	0.004464538	0.000480779
	Hispanic	0.001390021	0.008730844	0.000752169
22312	White	0.018770807	0.071728004	0.004991073
	African American	0.005922762	0.03534132	0.003805853
	Hispanic	0.005490728	0.034487747	0.002971146
22101	White	0.032716163	0.125016743	0.00869908
	African American	0.000458441	0.002735534	0.000294586
	Hispanic	0.000999449	0.006277628	0.000540823
20170	White	0.025286562	0.096626354	0.006723582
	African American	0.002261409	0.013493902	0.001453138
	Hispanic	0.004767609	0.029945776	0.002579852
20194	White	0.032401325	0.123813664	0.008615366
	African American	0.001478965	0.008825035	0.000950354
	Hispanic	0.001096599	0.006887832	0.000593392
20190	White	0.027723014	0.105936656	0.007371424
	African American	0.002701172	0.016117984	0.001735721
	Hispanic	0.003087215	0.01939107	0.001670556
22182	White	0.030710885	0.117354065	0.008165885
	African American	0.00078115	0.004661148	0.000501952
	Hispanic	0.000899843	0.005651991	0.000486923
22043	White	0.027071606	0.103447462	0.007198218
	African American	0.001357052	0.008097578	0.000872016
	Hispanic	0.003560967	0.022366747	0.001926913
22046	White	0.032038579	0.122427523	0.008518913
	African American	0.000880655	0.005254897	0.000565892
	Hispanic	0.002764257	0.017362539	0.001495796

22044	White	0.024064671	0.091957199	0.006398687
	African American	0.001398233	0.008343302	0.000898477
	Hispanic	0.007323295	0.045998261	0.003962786
22041	White	0.019882519	0.075976139	0.005286672
	African American	0.002685009	0.016021537	0.001725335
	Hispanic	0.008619934	0.054142568	0.004664424
20171	White	0.029180106	0.111504571	0.007758858
	African American	0.001424442	0.008499691	0.000915319
	Hispanic	0.001428734	0.008974004	0.000773117
20191	White	0.027262003	0.104175017	0.007248843
	African American	0.002815927	0.016802729	0.00180946
	Hispanic	0.003064996	0.019251508	0.001658532
22181	White	0.032046397	0.122457396	0.008520992
	African American	0.000743729	0.004437856	0.000477906
	Hispanic	0.00166266	0.010443314	0.0008997
22180	White	0.030047456	0.114818936	0.007989483
	African American	0.00093122	0.005556619	0.000598384
	Hispanic	0.002313098	0.014528775	0.001251665

22027	White	0.029093221	0.111172564	0.007735756
	African American	0.000819747	0.00489146	0.000526754
	Hispanic	0.000797302	0.005007923	0.000431436
22042	White	0.024033646	0.091838646	0.006390438
	African American	0.001307457	0.007801642	0.000840147
	Hispanic	0.005687307	0.035722477	0.003077519
22312	White	0.018770807	0.071728004	0.004991073
	African American	0.005922762	0.03534132	0.003805853
	Hispanic	0.005490728	0.034487747	0.002971146
22310	White	0.027519099	0.105157445	0.007317204
	African American	0.002979009	0.017775848	0.001914254
	Hispanic	0.002508973	0.015759079	0.001357657
22303	White	0.025707003	0.098232967	0.006835376
	African American	0.003827588	0.02283935	0.002459535
	Hispanic	0.004323367	0.027155447	0.002339463
22307	White	0.032910447	0.125759148	0.008750739
	African American	0.00185771	0.011085017	0.001193729
	Hispanic	0.001912565	0.012012987	0.001034928
22308	White	0.03609881	0.13794269	0.00959851
	African American	0.000672221	0.004011166	0.000431956
	Hispanic	0.000658948	0.00413891	0.00035657
22306	White	0.019495446	0.074497036	0.005183751
	African American	0.007329559	0.043735724	0.004709833
	Hispanic	0.004944836	0.03105895	0.002675753
22309	White	0.020898917	0.079860052	0.005556928
	African American	0.007606554	0.045388563	0.004887825
	Hispanic	0.003692086	0.023190316	0.001997864
22315	White	0.027801032	0.106234782	0.007392169
	African American	0.003564896	0.021271853	0.002290733
	Hispanic	0.001681085	0.01055904	0.00090967
22060	White	0.022395837	0.085580164	0.005954952
	African American	0.008192203	0.048883152	0.005264152
	Hispanic	0.002477569	0.01556183	0.001340664

22079	White	0.020376523	0.077863854	0.005418025
	African American	0.008627406	0.051480021	0.005543805
	Hispanic	0.002366094	0.014861646	0.001280342
22039	White	0.03270619	0.124978633	0.008696428
	African American	0.001001066	0.005973394	0.000643266
	Hispanic	0.000732611	0.004601597	0.000396431
22153	White	0.028126155	0.107477158	0.007478617
	African American	0.002591444	0.015463231	0.001665212
	Hispanic	0.001779258	0.011175677	0.000962793
22150	White	0.022384	0.085534931	0.005951804
	African American	0.001983509	0.011835664	0.001274565
	Hispanic	0.004763144	0.029917728	0.002577436
22152	White	0.028649363	0.109476469	0.007617736
	African American	0.001559105	0.009303232	0.001001851
	Hispanic	0.002100556	0.013193778	0.001136654
22151	White	0.027688178	0.105803537	0.007362161
	African American	0.00116824	0.006970927	0.000750688
	Hispanic	0.002456707	0.01543079	0.001329375
22003	White	0.025543122	0.097606735	0.006791801
	African American	0.001525519	0.009102825	0.000980269
	Hispanic	0.003454592	0.021698594	0.001869351
22015	White	0.028406733	0.108549317	0.007553222
	African American	0.001321329	0.007884416	0.000849061
	Hispanic	0.002136504	0.013419568	0.001156106
22032	White	0.030234756	0.115534656	0.008039285
	African American	0.001104404	0.006590016	0.000709669
	Hispanic	0.001408045	0.008844056	0.000761922
22030	White	0.027802254	0.106239453	0.007392494
	African American	0.002012071	0.012006093	0.001292918
	Hispanic	0.002468331	0.015503803	0.001335665
22031	White	0.024554371	0.093828466	0.006528896
	African American	0.001467899	0.008759001	0.000943243
	Hispanic	0.003527274	0.02215512	0.001908681

20124	White	0.032834812	0.12547013	0.008730628
	African American	0.000761273	0.00454254	0.000489179
	Hispanic	0.000877814	0.005513629	0.000475003
20121	White	0.025707792	0.098235983	0.006835586
	African American	0.002758497	0.016460047	0.001772557
	Hispanic	0.002596063	0.0163061	0.001404783
20120	White	0.029557579	0.112946993	0.007859227
	African American	0.001885198	0.011249042	0.001211392
	Hispanic	0.001698829	0.010670493	0.000919271
20151	White	0.028177439	0.107673129	0.007492254
	African American	0.001426642	0.008512823	0.000916733
	Hispanic	0.002655575	0.0166799	0.001436986
20171	White	0.029180106	0.111504571	0.007758858
	African American	0.001424442	0.008499691	0.000915319
	Hispanic	0.001428734	0.008974004	0.000773117
22033	White	0.029279248	0.111883418	0.00778522
	African American	0.001358319	0.008105133	0.000872829
	Hispanic	0.001392399	0.008745778	0.000753456
22124	White	0.033922931	0.129628107	0.009019954
	African American	0.000806261	0.004810984	0.000518088
	Hispanic	0.001029242	0.00646476	0.000556944
22304	White	0.021189204	0.080969312	0.005634113
	African American	0.006552356	0.03909813	0.004210418
	Hispanic	0.003820153	0.023994717	0.002067163

Table 15: Alcohol Tree Simulator - Detailed Output

LOW Availability	MEDIUM Availability	HIGH Availability
[outlet Avail: LOW AVAIL -race: BLACK -jobClass: WHITE COLLAR -alcohol Usage: MISUSER -number of Acute Outcomes = 8 -DWI = 7, -ASSAULT = 0, -MURDER = 0, -SEXUAL_ASSAULT = 0, -DOMESTIC_VIOLENCE = 0, -CHILD_ABUSE = 1, -SUICIDE = 0 -number of Benign Outcomes = 4006, -alcohol Usage: NONMISUSER -number of Acute Outcomes = 60 -DWI = 26, -ASSAULT = 9, -MURDER = 0, -SEXUAL_ASSAULT = 1, -DOMESTIC_VIOLENCE = 8, -CHILD_ABUSE = 10, -SUICIDE = 6 -number of Benign Outcomes = 54965, -jobClass: BLUE COLLAR -alcohol Usage: MISUSER -number of Acute Outcomes = 14 -DWI = 8, -ASSAULT = 0, -MURDER = 0, -SEXUAL_ASSAULT = 0, -DOMESTIC_VIOLENCE = 4, -CHILD_ABUSE = 1, -SUICIDE = 1 -number of Benign Outcomes = 9284, -alcohol Usage: NONMISUSER -number of Acute Outcomes = 176 -DWI = 91, -ASSAULT = 15, -MURDER = 2, -SEXUAL_ASSAULT = 3, -DOMESTIC_VIOLENCE = 24, -CHILD_ABUSE = 34, -SUICIDE = 7 -number of Benign Outcomes = 123300, -jobClass: UNEMPLOYED -alcohol Usage: MISUSER -number of Acute Outcomes = 1 -DWI = 1, -ASSAULT = 0, -MURDER = 0, -SEXUAL_ASSAULT = 0, -DOMESTIC_VIOLENCE = 0,	[outlet Avail: MED AVAIL -race: BLACK -jobClass: WHITE COLLAR -alcohol Usage: MISUSER -number of Acute Outcomes = 18 -DWI = 11, -ASSAULT = 1, -MURDER = 0, -SEXUAL_ASSAULT = 0, -DOMESTIC_VIOLENCE = 4, -CHILD_ABUSE = 0, -SUICIDE = 2 -number of Benign Outcomes = 8129, -alcohol Usage: NONMISUSER -number of Acute Outcomes = 170 -DWI = 84, -ASSAULT = 18, -MURDER = 1, -SEXUAL_ASSAULT = 3, -DOMESTIC_VIOLENCE = 22, -CHILD_ABUSE = 29, -SUICIDE = 13 -number of Benign Outcomes = 109826, -jobClass: BLUE COLLAR -alcohol Usage: MISUSER -number of Acute Outcomes = 22 -DWI = 16, -ASSAULT = 1, -MURDER = 1, -SEXUAL_ASSAULT = 0, -DOMESTIC_VIOLENCE = 2, -CHILD_ABUSE = 1, -SUICIDE = 1 -number of Benign Outcomes = 18354, -alcohol Usage: NONMISUSER -number of Acute Outcomes = 326 -DWI = 161, -ASSAULT = 45, -MURDER = 2, -SEXUAL_ASSAULT = 7, -DOMESTIC_VIOLENCE = 38, -CHILD_ABUSE = 46, -SUICIDE = 27 -number of Benign Outcomes = 246301, -jobClass: UNEMPLOYED -alcohol Usage: MISUSER -number of Acute Outcomes = 5 -DWI = 3, -ASSAULT = 0, -MURDER = 0, -SEXUAL_ASSAULT = 0, -DOMESTIC_VIOLENCE = 0,	[outlet Avail: HIGH AVAIL -race: BLACK -jobClass: WHITE COLLAR -alcohol Usage: MISUSER -number of Acute Outcomes = 2 -DWI = 0, -ASSAULT = 1, -MURDER = 0, -SEXUAL_ASSAULT = 0, -DOMESTIC_VIOLENCE = 0, -CHILD_ABUSE = 1, -SUICIDE = 0 -number of Benign Outcomes = 1311, -alcohol Usage: NONMISUSER -number of Acute Outcomes = 28 -DWI = 15, -ASSAULT = 4, -MURDER = 0, -SEXUAL_ASSAULT = 0, -DOMESTIC_VIOLENCE = 3, -CHILD_ABUSE = 5, -SUICIDE = 1 -number of Benign Outcomes = 18375, -jobClass: BLUE COLLAR -alcohol Usage: MISUSER -number of Acute Outcomes = 3 -DWI = 1, -ASSAULT = 1, -MURDER = 0, -SEXUAL_ASSAULT = 0, -DOMESTIC_VIOLENCE = 1, -CHILD_ABUSE = 0, -SUICIDE = 0 -number of Benign Outcomes = 3084, -alcohol Usage: NONMISUSER -number of Acute Outcomes = 68 -DWI = 32, -ASSAULT = 11, -MURDER = 1, -SEXUAL_ASSAULT = 3, -DOMESTIC_VIOLENCE = 10, -CHILD_ABUSE = 6, -SUICIDE = 5 -number of Benign Outcomes = 41022, -jobClass: UNEMPLOYED -alcohol Usage: MISUSER -number of Acute Outcomes = 0 -DWI = 0, -ASSAULT = 0, -MURDER = 0, -SEXUAL_ASSAULT = 0, -DOMESTIC_VIOLENCE = 0,

```
    -CHILD_ABUSE = 0,
    -SUICIDE = 0
   -number of Benign Outcomes = 914,
  -alcohol Usage: NONMISUSER
   -number of Acute Outcomes = 16
    -DWI = 6,
    -ASSAULT = 3,
    -MURDER = 0,
    -SEXUAL_ASSAULT = 0,
    -DOMESTIC_VIOLENCE = 2,
    -CHILD_ABUSE = 5,
    -SUICIDE = 0
   -number of Benign Outcomes = 11528,
 -race: WHITE
  -jobClass: WHITE COLLAR
  -alcohol Usage: MISUSER
   -number of Acute Outcomes = 0
    -DWI = 0,
    -ASSAULT = 0,
    -MURDER = 0,
    -SEXUAL_ASSAULT = 0,
    -DOMESTIC_VIOLENCE = 0,
    -CHILD_ABUSE = 0,
    -SUICIDE = 0
   -number of Benign Outcomes = 538,
  -alcohol Usage: NONMISUSER
   -number of Acute Outcomes = 7
    -DWI = 4,
    -ASSAULT = 0,
    -MURDER = 0,
    -SEXUAL_ASSAULT = 0,
    -DOMESTIC_VIOLENCE = 2,
    -CHILD_ABUSE = 0,
    -SUICIDE = 1
   -number of Benign Outcomes = 7230,
  -jobClass: BLUE COLLAR
  -alcohol Usage: MISUSER
   -number of Acute Outcomes = 4
    -DWI = 2,
    -ASSAULT = 1,
    -MURDER = 0,
    -SEXUAL_ASSAULT = 0,
    -DOMESTIC_VIOLENCE = 1,
    -CHILD_ABUSE = 0,
    -SUICIDE = 0
   -number of Benign Outcomes = 1884,
  -alcohol Usage: NONMISUSER
   -number of Acute Outcomes = 40
    -DWI = 21,
    -ASSAULT = 1,
    -MURDER = 1,
    -SEXUAL_ASSAULT = 1,
    -DOMESTIC_VIOLENCE = 7,
    -CHILD_ABUSE = 6,
    -SUICIDE = 3
   -number of Benign Outcomes = 25777,
  -jobClass: UNEMPLOYED
  -alcohol Usage: MISUSER
   -number of Acute Outcomes = 0
    -DWI = 0,
    -ASSAULT = 0,
```

```
    -CHILD_ABUSE = 1,
    -SUICIDE = 1
   -number of Benign Outcomes = 1657,
  -alcohol Usage: NONMISUSER
   -number of Acute Outcomes = 24
    -DWI = 11,
    -ASSAULT = 3,
    -MURDER = 0,
    -SEXUAL_ASSAULT = 0,
    -DOMESTIC_VIOLENCE = 3,
    -CHILD_ABUSE = 6,
    -SUICIDE = 1
   -number of Benign Outcomes = 22596,
 -race: WHITE
  -jobClass: WHITE COLLAR
  -alcohol Usage: MISUSER
   -number of Acute Outcomes = 0
    -DWI = 0,
    -ASSAULT = 0,
    -MURDER = 0,
    -SEXUAL_ASSAULT = 0,
    -DOMESTIC_VIOLENCE = 0,
    -CHILD_ABUSE = 0,
    -SUICIDE = 0
   -number of Benign Outcomes = 1092,
  -alcohol Usage: NONMISUSER
   -number of Acute Outcomes = 17
    -DWI = 9,
    -ASSAULT = 1,
    -MURDER = 0,
    -SEXUAL_ASSAULT = 1,
    -DOMESTIC_VIOLENCE = 3,
    -CHILD_ABUSE = 2,
    -SUICIDE = 1
   -number of Benign Outcomes = 14595,
  -jobClass: BLUE COLLAR
  -alcohol Usage: MISUSER
   -number of Acute Outcomes = 8
    -DWI = 5,
    -ASSAULT = 1,
    -MURDER = 0,
    -SEXUAL_ASSAULT = 0,
    -DOMESTIC_VIOLENCE = 1,
    -CHILD_ABUSE = 0,
    -SUICIDE = 1
   -number of Benign Outcomes = 3759,
  -alcohol Usage: NONMISUSER
   -number of Acute Outcomes = 80
    -DWI = 39,
    -ASSAULT = 9,
    -MURDER = 0,
    -SEXUAL_ASSAULT = 4,
    -DOMESTIC_VIOLENCE = 13,
    -CHILD_ABUSE = 10,
    -SUICIDE = 5
   -number of Benign Outcomes = 50466,
  -jobClass: UNEMPLOYED
  -alcohol Usage: MISUSER
   -number of Acute Outcomes = 1
```

```
    -CHILD_ABUSE = 0,
    -SUICIDE = 0
   -number of Benign Outcomes = 272,
  -alcohol Usage: NONMISUSER
   -number of Acute Outcomes = 6
    -DWI = 6,
    -ASSAULT = 0,
    -MURDER = 0,
    -SEXUAL_ASSAULT = 0,
    -DOMESTIC_VIOLENCE = 0,
    -CHILD_ABUSE = 0,
    -SUICIDE = 0
   -number of Benign Outcomes = 3838,
 -race: WHITE
  -jobClass: WHITE COLLAR
  -alcohol Usage: MISUSER
   -number of Acute Outcomes = 0
    -DWI = 0,
    -ASSAULT = 0,
    -MURDER = 0,
    -SEXUAL_ASSAULT = 0,
    -DOMESTIC_VIOLENCE = 0,
    -CHILD_ABUSE = 0,
    -SUICIDE = 0
   -number of Benign Outcomes = 174,
  -alcohol Usage: NONMISUSER
   -number of Acute Outcomes = 4
    -DWI = 3,
    -ASSAULT = 0,
    -MURDER = 0,
    -SEXUAL_ASSAULT = 0,
    -DOMESTIC_VIOLENCE = 0,
    -CHILD_ABUSE = 1,
    -SUICIDE = 0
   -number of Benign Outcomes = 2505,
  -jobClass: BLUE COLLAR
  -alcohol Usage: MISUSER
   -number of Acute Outcomes = 1
    -DWI = 1,
    -ASSAULT = 0,
    -MURDER = 0,
    -SEXUAL_ASSAULT = 0,
    -DOMESTIC_VIOLENCE = 0,
    -CHILD_ABUSE = 0,
    -SUICIDE = 0
   -number of Benign Outcomes = 561,
  -alcohol Usage: NONMISUSER
   -number of Acute Outcomes = 14
    -DWI = 2,
    -ASSAULT = 1,
    -MURDER = 0,
    -SEXUAL_ASSAULT = 0,
    -DOMESTIC_VIOLENCE = 3,
    -CHILD_ABUSE = 6,
    -SUICIDE = 2
   -number of Benign Outcomes = 8435,
  -jobClass: UNEMPLOYED
  -alcohol Usage: MISUSER
   -number of Acute Outcomes = 0
```

-MURDER = 0,
-SEXUAL_ASSAULT = 0,
-DOMESTIC_VIOLENCE = 0,
-CHILD_ABUSE = 0,
-SUICIDE = 0
-number of Benign Outcomes = 276,
-alcohol Usage: NONMISUSER
-number of Acute Outcomes = 2
-DWI = 0,
-ASSAULT = 0,
-MURDER = 0,
-SEXUAL_ASSAULT = 0,
-DOMESTIC_VIOLENCE = 0,
-CHILD_ABUSE = 2,
-SUICIDE = 0
-number of Benign Outcomes = 3590,
-race: HISPANIC
-jobClass: WHITE COLLAR
-alcohol Usage: MISUSER
-number of Acute Outcomes = 3
-DWI = 2,
-ASSAULT = 1,
-MURDER = 0,
-SEXUAL_ASSAULT = 0,
-DOMESTIC_VIOLENCE = 0,
-CHILD_ABUSE = 0,
-SUICIDE = 0
-number of Benign Outcomes = 585,
-alcohol Usage: NONMISUSER
-number of Acute Outcomes = 9
-DWI = 5,
-ASSAULT = 2,
-MURDER = 0,
-SEXUAL_ASSAULT = 1,
-DOMESTIC_VIOLENCE = 0,
-CHILD_ABUSE = 1,
-SUICIDE = 0
-number of Benign Outcomes = 7432,
-jobClass: BLUE COLLAR
-alcohol Usage: MISUSER
-number of Acute Outcomes = 1
-DWI = 1,
-ASSAULT = 0,
-MURDER = 0,
-SEXUAL_ASSAULT = 0,
-DOMESTIC_VIOLENCE = 0,
-CHILD_ABUSE = 0,
-SUICIDE = 0
-number of Benign Outcomes = 2134,
-alcohol Usage: NONMISUSER
-number of Acute Outcomes = 49
-DWI = 24,
-ASSAULT = 4,
-MURDER = 0,
-SEXUAL_ASSAULT = 1,
-DOMESTIC_VIOLENCE = 9,
-CHILD_ABUSE = 4,
-SUICIDE = 7
-number of Benign Outcomes = 28315,
-jobClass: UNEMPLOYED
-alcohol Usage: MISUSER

-DWI = 1,
-ASSAULT = 0,
-MURDER = 0,
-SEXUAL_ASSAULT = 0,
-DOMESTIC_VIOLENCE = 0,
-CHILD_ABUSE = 0,
-SUICIDE = 0
-number of Benign Outcomes = 533,
-alcohol Usage: NONMISUSER
-number of Acute Outcomes = 11
-DWI = 6,
-ASSAULT = 0,
-MURDER = 0,
-SEXUAL_ASSAULT = 0,
-DOMESTIC_VIOLENCE = 3,
-CHILD_ABUSE = 2,
-SUICIDE = 0
-number of Benign Outcomes = 7200,
-race: HISPANIC
-jobClass: WHITE COLLAR
-alcohol Usage: MISUSER
-number of Acute Outcomes = 4
-DWI = 2,
-ASSAULT = 1,
-MURDER = 0,
-SEXUAL_ASSAULT = 0,
-DOMESTIC_VIOLENCE = 0,
-CHILD_ABUSE = 1,
-SUICIDE = 0
-number of Benign Outcomes = 1142,
-alcohol Usage: NONMISUSER
-number of Acute Outcomes = 16
-DWI = 11,
-ASSAULT = 0,
-MURDER = 0,
-SEXUAL_ASSAULT = 1,
-DOMESTIC_VIOLENCE = 0,
-CHILD_ABUSE = 2,
-SUICIDE = 2
-number of Benign Outcomes = 14890,
-jobClass: BLUE COLLAR
-alcohol Usage: MISUSER
-number of Acute Outcomes = 5
-DWI = 4,
-ASSAULT = 0,
-MURDER = 0,
-SEXUAL_ASSAULT = 0,
-DOMESTIC_VIOLENCE = 0,
-CHILD_ABUSE = 0,
-SUICIDE = 1
-number of Benign Outcomes = 4359,
-alcohol Usage: NONMISUSER
-number of Acute Outcomes = 99
-DWI = 57,
-ASSAULT = 6,
-MURDER = 0,
-SEXUAL_ASSAULT = 1,
-DOMESTIC_VIOLENCE = 9,
-CHILD_ABUSE = 21,
-SUICIDE = 5
-number of Benign Outcomes =

-DWI = 0,
-ASSAULT = 0,
-MURDER = 0,
-SEXUAL_ASSAULT = 0,
-DOMESTIC_VIOLENCE = 0,
-CHILD_ABUSE = 0,
-SUICIDE = 0
-number of Benign Outcomes = 83,
-alcohol Usage: NONMISUSER
-number of Acute Outcomes = 1
-DWI = 0,
-ASSAULT = 0,
-MURDER = 0,
-SEXUAL_ASSAULT = 0,
-DOMESTIC_VIOLENCE = 1,
-CHILD_ABUSE = 0,
-SUICIDE = 0
-number of Benign Outcomes = 1188,
-race: HISPANIC
-jobClass: WHITE COLLAR
-alcohol Usage: MISUSER
-number of Acute Outcomes = 0
-DWI = 0,
-ASSAULT = 0,
-MURDER = 0,
-SEXUAL_ASSAULT = 0,
-DOMESTIC_VIOLENCE = 0,
-CHILD_ABUSE = 0,
-SUICIDE = 0
-number of Benign Outcomes = 188,
-alcohol Usage: NONMISUSER
-number of Acute Outcomes = 2
-DWI = 1,
-ASSAULT = 0,
-MURDER = 0,
-SEXUAL_ASSAULT = 0,
-DOMESTIC_VIOLENCE = 0,
-CHILD_ABUSE = 0,
-SUICIDE = 1
-number of Benign Outcomes = 2406,
-jobClass: BLUE COLLAR
-alcohol Usage: MISUSER
-number of Acute Outcomes = 0
-DWI = 0,
-ASSAULT = 0,
-MURDER = 0,
-SEXUAL_ASSAULT = 0,
-DOMESTIC_VIOLENCE = 0,
-CHILD_ABUSE = 0,
-SUICIDE = 0
-number of Benign Outcomes = 740,
-alcohol Usage: NONMISUSER
-number of Acute Outcomes = 15
-DWI = 7,
-ASSAULT = 0,
-MURDER = 0,
-SEXUAL_ASSAULT = 1,
-DOMESTIC_VIOLENCE = 0,
-CHILD_ABUSE = 5,
-SUICIDE = 2
-number of Benign Outcomes =

-number of Acute Outcomes = 0 -DWI = 0, -ASSAULT = 0, -MURDER = 0, -SEXUAL_ASSAULT = 0, -DOMESTIC_VIOLENCE = 0, -CHILD_ABUSE = 0, -SUICIDE = 0 -number of Benign Outcomes = 210, -alcohol Usage: NONMISUSER -number of Acute Outcomes = 2 -DWI = 0, -ASSAULT = 1, -MURDER = 0, -SEXUAL_ASSAULT = 0, -DOMESTIC_VIOLENCE = 1, -CHILD_ABUSE = 0, -SUICIDE = 0 -number of Benign Outcomes = 3077]	57277, -jobClass: UNEMPLOYED -alcohol Usage: MISUSER -number of Acute Outcomes = 0 -DWI = 0, -ASSAULT = 0, -MURDER = 0, -SEXUAL_ASSAULT = 0, -DOMESTIC_VIOLENCE = 0, -CHILD_ABUSE = 0, -SUICIDE = 0 -number of Benign Outcomes = 457, -alcohol Usage: NONMISUSER -number of Acute Outcomes = 8 -DWI = 3, -ASSAULT = 0, -MURDER = 0, -SEXUAL_ASSAULT = 0, -DOMESTIC_VIOLENCE = 3, -CHILD_ABUSE = 2, -SUICIDE = 0 -number of Benign Outcomes = 6156]	9521, -jobClass: UNEMPLOYED -alcohol Usage: MISUSER -number of Acute Outcomes = 0 -DWI = 0, -ASSAULT = 0, -MURDER = 0, -SEXUAL_ASSAULT = 0, -DOMESTIC_VIOLENCE = 0, -CHILD_ABUSE = 0, -SUICIDE = 0 -number of Benign Outcomes = 78, -alcohol Usage: NONMISUSER -number of Acute Outcomes = 0 -DWI = 0, -ASSAULT = 0, -MURDER = 0, -SEXUAL_ASSAULT = 0, -DOMESTIC_VIOLENCE = 0, -CHILD_ABUSE = 0, -SUICIDE = 0 -number of Benign Outcomes = 1024]

Curriculum Vitae

Yasmin Said was born in Arecibo, Puerto Rico and was resident there until age five. She began elementary school in Tulkarem, Palestine (West Bank) from first through third grade. She came to the United States to finish her education. By the time she finished her elementary education, Yasmin was trilingual in Spanish, Arabic and English. She attended Rockville High School, where she was a journalist for her high school newspaper, an editor of the yearbook, and a photographer for both the newspaper and the yearbook. She graduated with an honors certificate of merit. She attended Trinity College where she earned an A.B. in pure mathematics with minor in bioethics and philosophy. She was awarded the Trinity College Leadership Award for four years, was on the National Dean's List, and the Trinity College Dean's List. She received an M.S. in Computer and Information Systems at American University in Washington, DC. At American University, she placed first in the regional mathematics competition, was awarded the American University Leadership Award and the AU Women's Leadership Award.

Subsequent to her graduation from Trinity College, she began a five-year high school teaching career in Washington, DC and Northern Virginia, teaching calculus, statistics, algebra, and finite mathematics. She enrolled in the Ph.D. Program in Computational Sciences in Fall 2003 and pursued the Ph.D. with a specialization in Computational Statistics with a 4.0 grade point average. Her dissertation, *Agent Based Simulation of Ecological Alcohol Systems*, introduces a new framework designed to simultaneously mitigate the acute effects of alcohol abuse. Her research can be viewed at the website http://www.alcoholecology.com.

During her time at George Mason University, she has published a research paper entitled "On Genetic Algorithms and their Applications" in the *Handbook of Statistics: Data Mining and Data Visualization*. She has organized sessions at both the Symposium on the Interface of Computing Science and Statistics and the Joint Statistical Meetings. She is a fellow of the Royal Statistical Society, an elected member of the Research Society on Alcoholism, and a member of the American Statistical Association, Institute of Mathematical Statistics, Institute of Electrical and Electronic Engineers, American Association for the Advancement of Science, American Mathematical Society, Society for Industrial and Applied Mathematics, and The Interface Foundation of North America.

Yasmin is planning an academic career focusing on teaching and research. Her immediate research goals are to expand and develop her dissertation research with a view to expanding her model from a local scale to a national scale and creating a more in-depth social network analysis of the complete alcohol ecology system including modeling the long-term effects of interventions.