

Probability Sampling Exercises

Instructions: (1) Define the theoretical population. (2) Define an accessible population that you could use for the study and discuss the strengths and weaknesses of the choice in terms of how well the accessible population represents the theoretical population. If I chose the accessible population for you, simply discuss the strengths and weaknesses. (3) Identify a sampling approach (e.g., stratified random, cluster, etc.) that you could use to achieve the greatest internal and external validity and the greatest efficiency in sampling for your study, making sure that you justify your decisions. (4) How would you determine the sample size? You don't have to actually calculate the sample size, but you must explain the factors you **would** include in making your decision. Specifically decide the confidence interval you would use. Explain how you would determine an appropriate level of precision. Explain how you would estimate the variance. Justify your choices. (5) Explain the steps that you would take to recruit the participants – e.g., how you would go about actually “finding” them and getting them to participate in your study.

Problem 1: Drinking Habits of College Students

You want to conduct a study of any change in drinking habits that may occur among undergraduate students during the period of time for completing the undergraduate degree. You specifically want to understand any patterns that develop as incoming freshmen complete their bachelor degree programs. You will limit your study to full time students. Therefore, you need to track **the same individuals** over the entire time they are enrolled as full-time undergraduate students – or set some cut-off that is justifiable. In addition, you want to know whether there are differences in these patterns due to race, ethnicity or gender. You will start your study in fall semester, 2007. You have decided to use the University of Florida incoming freshman class as your accessible population. The table below shows the race, gender and ethnicity of incoming freshmen – defined as first-time college students starting the first year of study. Remember that **attrition is a problem in longitudinal studies**. Therefore, you want to decide how to calculate the **final sample size(s)** you will need; e.g., the initial sample size(s) must be adjusted for attrition. Attrition refers to both those students who leave UF and to those with whom you lose contact, don't want to respond to you any more, etc. Explain the assumptions you make. Historically, 55% of incoming freshmen graduate in four years or less; 22% graduate in more than four years, but in five years or less; 4% graduate in more than five years, but in six year or less. The overall graduation rate for six years or less of study is 79%.

Table 1.1 Freshmen admitted at the University of Florida, fall, 2007, by race/ethnicity and gender

Race/Ethnicity	Female	Male	Total
Nonresident aliens	17	18	35
Black, non-Hispanic	565	331	896
American Indian or Alaskan Native	11	9	20
Asian or Pacific Islander	282	271	552
White, non-Hispanic	2112	1872	3984
Not reported	137	143	280
Total	3124	2644	6702

Problem 2: Environmental Impacts of Nurseries

Potted plant nursery production (anything from trees to poinsettias, but not things like sod) is a major agricultural industry in Florida. This is an intensive form of agricultural production that can have a big environmental impact. You want to conduct a study of the degree to which Florida's nursery operators have adopted several agricultural production practices and technologies, such as recycling irrigation water, precision fertilizer application, using renewable resources as the potting media, and quarantine of imported plants, that can protect the environment from unwanted environmental impacts. Recycling irrigation water, for example, can reduce total water use and help make sure that fertilizer or pesticide residues in irrigation water do not enter streams or the aquifer. Your hypothesis is that nursery size is **negatively associated** with adoption of these technologies; e.g., the smaller nurseries will be less apt to use the desired practices and technologies. You also want to be able to draw conclusions about the potential impacts (positive or negative) of nursery production on the environment, things like total fresh water usage by Florida's nurseries. Obviously, the potential impact of production practices on the environment is **positively associated** with nursery size; e.g., a large nursery that recycles irrigation water will have a much greater impact on water conservation than a small one.

You will use a multiple-comparison group cross-sectional design. The three groups will be small, medium and large nurseries. Industry experts tell you that nurseries that produce fewer than 75,000 plants per year are considered small, those that produce 75,000 to 500,000 plants are considered medium sized, and nurseries that produce more than 500,000 plants per year are considered large. In addition, experts say that nurseries that produce fewer than 5,000 plants per year are basically "hobby" nurseries that are very unstable. Many come into existence each year, but many also cease operation each year. Anyone who sells nursery plants in Florida must register with the Florida Department of Agriculture and Consumers Sciences (FDACS) and secure a sanitation permit certifying that the nursery does not harbor things like exotic pests. FDACS has given you the 2006 list, which includes telephone number, e-mail address, and physical location (address). FDACS divides nurseries into ten size classes, based on number of plants sold per year (see Table 2-1). The list is in alphabetical order (name of nursery), by size class. You will conduct personal interviews with the nursery operators and observe certain features of the operation (to verify the information provided in the interview). Your data will, for the most part, be nominal (use or does not use a practice) or ordinal (five categories indicating intensity of use of a practice or technology, like never, rarely, sometimes, often, and always). Every county in Florida has at least a few nurseries on the FDACS list, but 24 counties are home to 40% of all nurseries (see Table 2-2).

Table 2-1. Number of nurseries, by size class

Size Class	Plants Sold/Year	No. of Nurseries
Class 1	<1,000	79,432
Class 2	1,001-5,000	46,544
Class 3	5,001-10,000	23,414
Class 4	10,001-25,000	18,678
Class 5	25,001-50,000	11,551
Class 6	50,001-100,000	6,764
Class 7	1,001-500,000	3,467
Class 8	500,001-1,000,000	987
Class 9	1,001,001-5,000,000	643
Class 10	>5,000,000	76

Table 2-2. Number of nurseries in 24 Florida counties by region of the state

County	No. of Nurseries	Region of State
Gulf	2,712	Northwest
Leon	3,532	
Santa Rosa	1,587	
Duval	2,488	Northeast
Escambia	2,376	
Flagler	3,001	
Nassau	3,966	
St. Johns	2,481	
Citrus	2,942	Central
Hernando	3,507	
Orange	4,487	
Polk	4,844	
Seminole	4,212	
Sumter	4,222	
Volusia	5,304	
Collier	4,534	South
Dade	2,553	
Glades	4,915	
Hillsborough	2,163	
Lee	6,118	
Manatee	3,541	
Martin	4,686	
Palm Beach	3,282	
St. Lucie	3,021	

Problem 3: Unreported Crime

Under-reporting of crime, especially minor crimes, is a common phenomenon. Many people do not report crimes like vandalism, purse-snatching, minor theft or nuisances (noise, for example). Some do not report more serious crimes either. You have two alternative hypotheses to explain this phenomenon. One is that under-reporting grows out of a sense of frustration, the belief on the part of some citizens or groups of citizens that the police cannot or will not respond to crime reports. You hypothesize that under-reporting due to frustration will be **positively associated** with areas of high reported crimes. An alternative explanation is that under-reporting is associated with socio-economic characteristics of the victim. Your hypothesis is that under-reporting will be **negatively associated** with socio-economic status (e.g., individuals of higher socio-economic status are more apt to report crimes than individuals of lower socio-economic status). You want to conduct a study to determine which of these two explanations better explains unreported crime in Gainesville, Florida. You need a probability sample, but there is, obviously, no sampling frame for individuals who failed to report crimes. However, the Gainesville Police Department does provide crime statistics for **reported crimes** by district and zone (see Table 3-1 and Figure 3-1).

Table 3-1. Reported crimes by district and zone, Gainesville, Florida, 2006

District 1		District 2		District 3	
Zone	# of Crimes	Zone	# of Crimes	Zone	# of Crimes
Alpha	570	Hotel	1779	Foxtrot	661
Bravo	617	Lima	981	Golf	1482
Charlie	1893	Mike	1117	India	1299
Delta	1870	Oscar	1163	November	1915
Echo	935	Papa	1277	Whiskey	1026
		Tango	1892	Xray	639
		Uniform	1687		

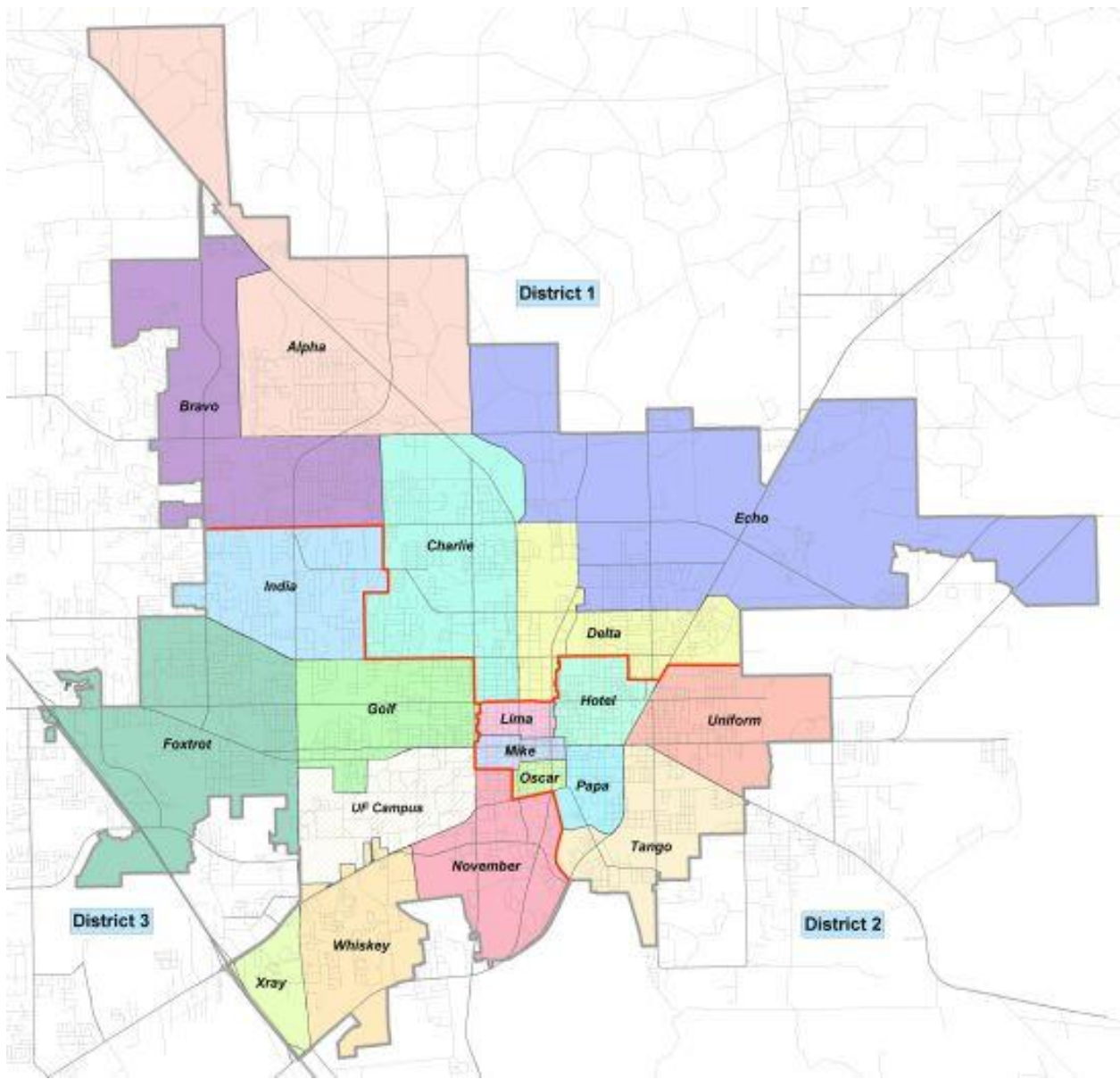


Figure 3-1. Gainesville Police Department districts and zones

Problem 4: Religiosity, Religious Affiliation and Attitude about Same-Sex Marriage

You want to examine the relationships between religiosity, religious affiliation, and attitudes or opinions about same-sex marriage. You have two hypotheses. The first is that religiosity will be **negatively associated** and directly associated with attitude or opinions about same sex marriage (e.g., greater religiosity = more negative opinions about same sex marriage). The second is that religious affiliation (faith and denomination) will influence religiosity directly and will indirectly influence attitude and opinion about same sex marriage. The second hypothesis is two-tailed. You have a well-tested index of religiosity that produces interval data. You also have a well-tested Likert scale that measures attitude/opinion about same sex marriage. It, too, generates interval data. You will conduct your study in Marion County, Florida. The Chamber of Commerce has provided you with a list of all churches in Marion County in 2006. They provided no membership data, but the list is by denomination (Table 4-1).

Table 4-1. Number of churches in Marion County Florida, by denomination, 2006

Denomination	# of Churches	Denomination	# of Churches
African Methodist Episcopal	7	Interdenominational	1
Anglican	1	Jewish	2
Apostolic	5	Jewish Reform (UAHC-affiliated)	1
Assembly of God	5	Lutheran	7
Baptist	56	Non-denominational	2
Catholic	7	Presbyterian	6
Christian	3	Salvation Army	1
Church of Christ	6	Seventh Day Adventist	2
Church of God	7	Unitarian	1
Episcopal	6	United Church of Christ	1
Friends (Quakers)	1	United Methodist	9
Grade Bethren	1	Unity	1

Problem 5: Young Adults' Participation in the Electoral Process

Despite many attempts to register young adults as voters, and considerable success in some cases, the number of young adults age 18-24 who vote is often low. Both the Democratic & Republican Parties conducted voter registration drives targeting young adults and both parties tried hard to "get out the youth vote" for the 2004 presidential election. In comparison to 2000, the number of young people who voted in the 2004 election was up sharply, especially in battleground states. Overall, voter turnout among 18-24 year-old citizens eligible to vote was up 11% in 2004 compared to 2000. However, compared to the historic high in 1972, electoral participation by this group actually fell by 5%.

A post-election study based on a nationwide sample of college students aged 18 to 24 showed that 87.6% were registered to vote in 2004 (Niemi, R.G. & Hanmer, M.J. [2006]. *Voter registration and turnout among college students*. Paper presented at the annual meeting of the American Political Science Association, Philadelphia, PA, August 31-September 3, 2006.). The researchers found that four factors strongly affected the probability of voting among college students in 2004. They were: (1) gender with females being more likely to vote than males ($p=0.001$), (2) frequency of discussing politics ($p=0.01$), (3) strength of partisanship ($p=0.001$), personal or mail contact by a political party ($p=0.001$), and prior voting history ($p=0.001$). You

want to know whether these same factors affect the probability of voting among non-college students aged 18 to 24 and to compare your results to those of the study of college student electoral participation. You will use a retrospective cross-sectional design that includes only those individuals who were aged 18 to 24 at the time of the 2004 election and who were not college students at that time. The original study was based on a statistically representative, completely random sample of 1200 college students at 285 randomly selected four-year universities. The original sample was statistically representative in terms of gender, type of institution, residence, region of the nation, and major (Table 5-1).

Table 5.1 Characteristics of the college student sample for the 2004 study

Characteristic	% of Sample
Gender	
Men	43.8
Women	56.2
Type of Institution	
Private	23.3
Public	76.7
Residence	
On-campus	68.7
Off-campus, not with parents	23.3
Off-campus, with parents	8.0
Region	
Northeast	19.3
Midwest	35.6
South	33.4
West	11.8
Major	
Business	14.6
Social sciences	13.6
Education	9.7
Engineering	6.9
Other	6.6

Problem 6: Reducing Child Exposure to Lead

You are writing a proposal to the National Institutes of Health (NIH) to initiate an educational program about the risks posed by lead in the home, especially for children aged 5 and younger. You will use a quasi-experimental design with treatment (get the educational program) and control (no educational program) communities. Ultimately, you expect to see differences in the outcome variables between the treatment and control communities. This study will involve testing for BLL (blood lead level) among children 5 and younger at the beginning of the study (before you implement your educational program), two years after starting the study and at the end of the study (four years after starting). You will also test (1) how much parents or caregivers know about how to identify lead in the home environment, (2) the risks that the lead poses to children, (3) telltale signs of effects of elevated BLL on the child's physical and mental development, and (4) who to go to for help to get BLL tests for children and/or a physician's assessment of any potential damage to children from exposure to lead in the home environment. Your hypothesis is that the change in all five variables will be greater among

children and families in the treatment group than in the control group. The proposal must include a sampling plan. While the NIH grant will be a large one, you cannot afford to include all of Duval County in your study. The funds simply will not be sufficient. You will be conducting intensive educational programs in the communities you select for the intervention and, of course, your three episodes of data collection require physically visiting each selected household. You will also have to visit the comparison or control group communities three times. Therefore, you need to devise a sampling plan that targets high risk areas in the county and, within those areas, permits you to identify the specific areas where lead exposure in the home environment is a high potential risk. Previous work shows that the risk of lead in the home environment is higher in older homes, and among the poor. Of course, you also want to make sure that your sample focuses on areas with high concentrations of children aged 5 and younger. You will want to match the control group communities to the intervention or treatment communities for concentrations of children aged 5 and younger.

Table 6.1 Median household income* and number of children aged 5 and younger by zip code, Jacksonville, FL

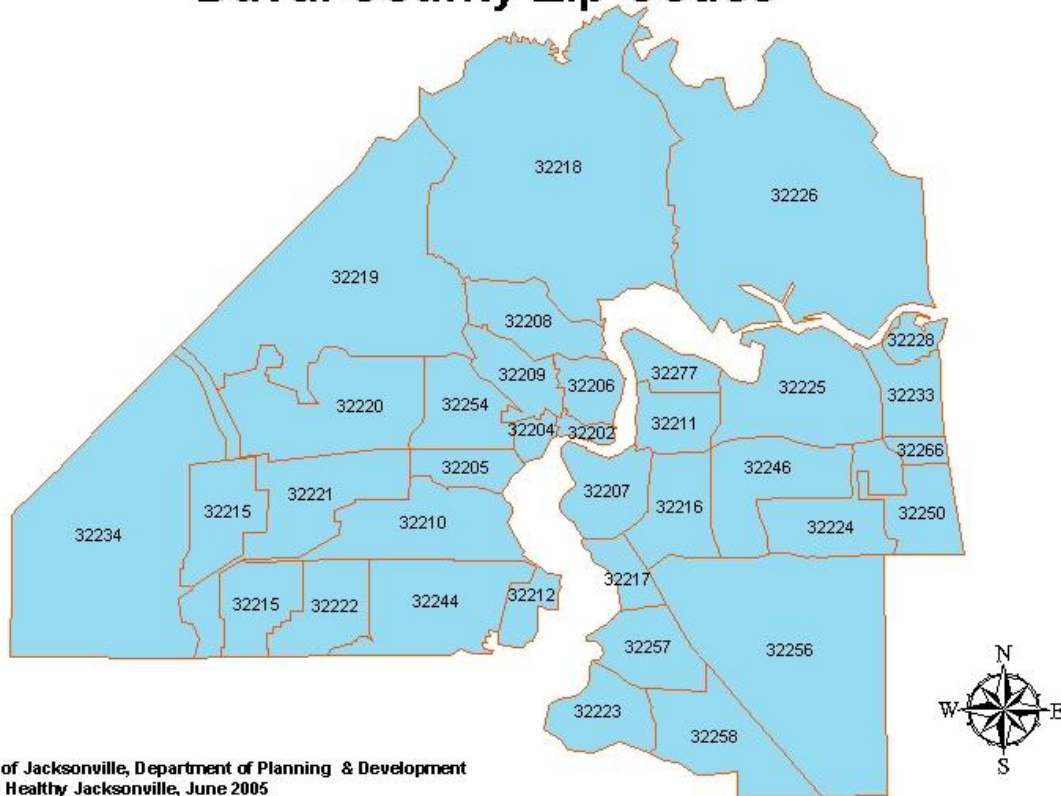
Zip Code	Median Household Income	Children 5 and Younger		Zip Code	Median Household Income	Children 5 and Younger	
		No.	% of all children			No.	% of all children
32202	9,904	171	67	32223	66,079	1,735	27
32204	23,957	470	33	32224	55,103	3,033	42
32205	34,579	2,109	34	32225	55,966	3,835	31
32206	17,169	1,377	28	32226	53,008	472	28
32207	35,568	2,473	35	32227	36,541	468	43
32208	31,849	1,575	23	32233	40,636	2,216	35
32209	22,202	2,420	28	32234	37,917	397	27
32210	39,095	4,488	31	32244	41,769	4,070	32
32211	34,232	2,741	34	32246	43,672	3,221	35
32212	34,653	282	40	32250	48,157	1,255	32
32215	40,714	208	43	32254	25,807	1,004	26
32216	41,354	2,018	30	32256	45,460	1,952	38
32217	39,740	1,194	27	32257	48,452	2,769	32
32218	40,691	2,619	29	32258	70,131	1,325	39
32219	36,961	586	29	32259	80,945	1,464	26
32220	41,622	749	27	32266	53,727	359	29
32221	42,816	1,209	29	32277	45,670	2,349	33
32222	43,641	378	35				

*National median \$41,994

Table 6-2. Number of housing units. by period of construction and zip code, Duval County, FL

Zip Code	Total	Period of Construction						
		1990-1999	1980-1989	1970-1979	1960-1969	1950-1959	1940-1949	1939 or earlier
32202	1,866	167	51	550	269	280	208	341
32204	3,910	52	171	461	656	588	529	1,453
32205	14,604	520	801	1,487	1,842	3,093	3,022	3,839
32206	10,464	229	598	793	1,503	1,946	1,973	3,422
32207	16,279	315	1,195	3,056	2,634	3,663	3,404	2,012
32208	13,746	500	577	1,811	3,987	3,995	2,008	868
32209	17,544	923	786	2,248	5,076	4,653	2,541	1,317
32210	24,598	1,900	3,990	5,364	4,793	5,192	2,253	1,106
32211	15,011	534	1,674	2,919	4,368	4,466	690	306
32212	337	6	41	72	73	105	40	0
32215	198	0	34	151	0	8	0	5
32216	12,476	795	2,451	2,782	2,937	2,578	712	221
32217	8,961	720	1,837	2,166	2,127	1,807	184	120
32218	14,781	3,665	3,054	2,520	2,470	2,184	642	246
32219	4,071	589	1,031	869	784	474	180	144
32220	4,181	1,184	1,130	927	435	313	157	35
32221	6,711	1,302	1,205	1,917	826	523	118	52
32222	1,768	710	534	276	87	110	41	10
32223	9,873	3,093	4,306	1,736	308	229	51	149
32224	13,507	10,022	2,251	856	255	70	31	22
32225	17,247	7,347	6,594	1,824	821	490	132	39
32226	3,185	1,023	876	556	353	254	64	59
32227	776	42	71	148	261	154	84	16
32233	10,684	2,192	3,657	2,059	1,412	881	279	204
32234	2,462	712	660	418	277	178	97	120
32244	18,606	5,358	6,135	3,267	2,004	1,323	414	105
32246	14,283	4,953	3,287	1,779	1,676	1,740	712	136
32250	12,042	2,576	2,508	2,532	1,488	1,808	683	447
32254	6,049	239	454	813	1,289	1,855	958	441
32256	15,244	6,086	5,429	2,981	473	146	120	9
32257	15,647	3,888	7,556	3,320	541	215	57	70
32258	4,380	2,175	1,633	414	94	51	7	6
32259	6,298	4,090	994	526	374	159	101	54
32266	3,459	341	790	680	718	367	247	316
32277	11,384	2,182	2,032	3,145	2,662	1,193	123	47

Duval County Zip Codes



Source: City of Jacksonville, Department of Planning & Development
Prepared by: Healthy Jacksonville, June 2005

This product is for reference only and is not to be construed as a legal document. Any reliance on the information contained herein is at the user's own risk. The Florida Department of Health and its agents assume no responsibility for any use of the information contained herein or any loss resulting therefrom.

Problem 7: Breast Cancer Screening Program

The incidence of breast cancer is no higher among African American women than among white women in the United States. However, the five-year survival rates for African American women are lower than they are for whites. Differences in survival are due in part to older African American women being diagnosed at a later stage of the disease and an ethnic gap in the use of mammography screening among women over 40 accounts for much of the difference. You conducted a short study of African American women who incurred breast cancer and found that the most common and most trusted source of information about women's health for these women was other women in their communities, rather than professional health care providers. You also found that women who had mammograms most commonly did so because another trusted woman in their community encouraged them to do so. Your study showed that these breast cancer victims were most likely to turn to other women who were slightly younger than they, active members of community organizations, and respected for their good judgment and willingness to help others. You now have a grant to recruit, train and implement a program of "lay health advisors" (LHAs) as a strategy for enhancing informational, emotional and instrumental support for mammography screening in five rural counties in Mississippi. You will recruit 30 women in each county who fit the description of a "trusted advisor" described above. These LHAs will participate in a three-day training program that provides them with knowledge and skills needed to help women over 40 overcome the reticence, lack of knowledge, and prevailing fears that leads them to neglect mammography screening. Each LHA is expected to

conduct both group sessions, as programs sponsored by community-based organizations, for example, and individual outreach to help reach the target audience. You have funds for the first three years of operation. Your donor for this program has indicated a willingness to extend funding for more counties and on a more permanent basis **IF** you can provide positive evidence that the intervention is effective in increasing the use of mammography screening among the target population. The donor's standards are high. You must provide strong statistical evidence that the self-reported use of mammography screening among African-American women over 40 is higher in the five counties where the program is implemented than in five other rural counties in Mississippi.

Problem 8: Long Term Reforestation Project

An international environmental organization has purchased 2010 hectares of degraded pastureland in Brazil. This land was originally part of the Amazon rainforest. In the mid-1980s 80 landless families were provided with loans to purchase of 25 hectares each to establish farms and the Brazilian government built a service center on 10 hectares. The original intent was that the smallholders would pay for the land over a 20-year period. However, after only a few years, crop production became virtually impossible without intensive application of agricultural production technology. Most of the smallholders then turned to cattle production as a lower cost alternative land use. In an effort to earn a living, most tried to raise too many cattle on too little land. As a result, land degradation was rapid and devastating. The environmental organization then offered to buy the land and eventually succeeded in getting all 80 landowners and the Brazilian government to agree to a lump sum settlement for all involved. The organization's interest was not primarily in preserving this degraded land, but rather in determining how (and *if*) the land would return to forest. Your job is to establish a prospective longitudinal study of the hypothesized reforestation process. The environmental organization is committed to maintaining the study for at least 25 years and perhaps longer if needed. It will obviously be impossible to monitor the entire 2010 hectares. Therefore, you need to demarcate specific small plots within the 2010 hectares that will be monitored yearly for species composition and growth rates of individual trees that establish in the plots. The site contains three ecologically different areas. (1) A third-order river with a wide floodplain that is subject to several weeks of partial or total inundation per year occupies the entire northern perimeter of the site. This area was originally covered by a forest complex dominated by tree species tolerant to inundation. The soils in the floodplain are young alluvial Inceptisols. (2) Adjacent to the floodplain there are a series of highly eroded terraces that reflect previous geologic periods in which the floodplain was more extensive than it is in recent geological history. The soils are Ultisols. Data from extant landscapes in the Amazon Basin are inconclusive about whether the forest complex that occupies these ancient terraces differs from that of other settings. (3) Most of the site consists of an ancient dissected upland originally occupied by a mature evergreen rainforest complex. The soils are Oxisols. Explain how you will decide where to establish the long-term observation plots, the number of plots, and sampling within the plots.

Problem 9: Physical Education in School and Childhood Obesity

Many have argued that removing physical education from the public school curriculum has directly contributed to the growing problem of obesity among children. They argue that even a few minutes per day of physical activity can have a major impact – especially for children whose home environments are conducive to obesity. Others argue that the role of a few minutes of physical activity during school is a minor effect, completely outweighed by the youth's home environment (eating habits, activity habits, etc.). An excellent opportunity to test the impact of physical education in the school on childhood obesity has arisen because the new Governor of

Florida, Gov. Christ, signed into law House Bill 967 which requires elementary schools to provide 150 minutes per week of Physical Education to students in grades K – 5. You want to conduct a five-year quasi-experiment that examines the impact of the new law on the incidence of childhood obesity, taking into account the level of physical activity outside the school. You have to decide how to establish the comparison groups. This is possible because some school districts did offer physical education in previous years, and it is possible to determine the level of physical activity of children outside the school. A 2005 study showed that 14% of all Florida school children are at risk for being overweight, and 12.4% are overweight. However, 43.8% are trying to lose weight. That same study showed that 27.3% of students attended physical education class daily, even though this was not a state requirement at that time. Almost half, 45.6%, attended physical education class one or more days during an average school week. Students were also asked about overall physical activity (at home or at school): 12.2% said they participated in no vigorous or moderate physical activity during the past seven days; 35.2% said that they did not participate in at least 20 minutes of vigorous physical activity on three or more of the past seven days and did not do at least 30 minutes of moderate physical activity on five or more of the past seven days; and 42.7% said that they watched three or more hours of TV per day on an average school day. Since your funds are not unlimited, you will also have to decide how to select a sample from Florida’s public schools. You have to decide determine the characteristics you will use to select the samples. There are 1,884 public elementary schools in Florida 67 school districts. Table 9.1 shows enrollment in Florida’s public schools in November, 2006, the most recent date for which data are available.

Table 9.1. Enrollment, by grade and ethnicity, in Florida public schools, November, 2006

Grade	White Non-Hispanic	Black Non-Hispanic	Hispanic	Asian/Pacific Islander	American Indian/Alaskan Native	Multiracial	Total Female	Total Male	Total Membership
PK	16,779	15,468	12,126	701	90	1,864	19,945	27,083	47,028
KG	90,804	47,414	50,730	4,358	615	9,292	97,171	106,042	203,213
1	92,963	47,216	51,430	4,297	584	8,803	98,338	106,955	205,293
2	91,500	45,953	49,681	4,318	649	8,506	97,102	103,505	200,607
3	94,692	49,263	51,296	4,650	640	7,953	100,337	108,157	208,494
4	92,585	43,486	47,088	4,453	582	6,999	95,947	99,246	195,193
5	95,539	45,698	47,786	4,436	611	6,630	97,803	102,897	200,700
6	94,968	43,893	43,926	4,189	590	5,937	93,350	100,153	193,503
7	100,605	50,421	48,358	4,340	626	5,584	101,689	108,245	209,934
8	101,711	48,843	47,495	4,454	619	5,190	101,596	106,716	208,312
9	119,155	60,647	54,916	4,932	726	4,880	116,952	128,304	245,256
10	105,611	49,695	48,258	4,911	586	3,494	104,938	107,617	212,555
11	95,897	42,205	39,593	4,902	511	2,824	93,630	92,302	185,932
12	83,769	34,418	32,216	4,481	457	2,202	80,978	76,565	157,543
TOTALS	1,276,578	624,620	624,899	59,422	7,886	80,158	1,299,776	1,373,787	2,673,563

Problem 10: Effects of Cranberry Juice Consumption on Higher Cognitive Functions of the Elderly

There is considerable evidence that consumption of red berries (grapes, raspberries, strawberries, cranberries, etc.) increases the level of certain chemicals in the human body, some of which have been shown to be related to increases in higher cognitive functions (short-term memory, logic, pattern recognition, among others). Ocean Spray wants to conduct a 10-week test of the effects of drinking two servings (12 oz.) of cranberry juice on the short-term memory, logic and pattern recognition among the elderly. They require that the study be completed with individuals over 70 years of age with no existing clinical condition related to cognitive function (such as Alzheimer's disease or dementia). Some previous research has also shown a statistically significant positive correlation between physical activity and higher cognitive function. Ocean Spray therefore also wants to include only those people who engage in at least 20 minutes of moderate physical activity (such as walking) on at least three days per week on average. However, the company wants to **exclude** individuals who engage in strenuous physical activity (such as intense physical labor or working out at a gym). Participants will be randomly assigned to treatment (grape juice) and control (grape-flavored drink of equivalent caloric value) groups. Each participant will be asked to consume two separate servings of grape juice (or control) each day (morning and evening), to provide a blood sample every two weeks, and to complete a set of tests of cognitive function each week (about ½ hour per week). Welch's will pay each participant \$200. It's going to be hard to find this sample, but you have a good idea. Lots of elderly people use malls as an exercise venue – generally for walking.

Problem 11: Estimating Transient Killer Whale Population Densities

You want to estimate the abundance of killer whales (*O. orca* (Linnaeus, 1758) within the known haulout range of the western stock of Steller sea lions in US waters. This stock is listed as 'endangered', pursuant to the US Endangered Species Act of 1973 [16 United States Code, pp. 1,531–1,543 (Supp. IV 1974)] as amended. In response to the possible impact of predation on Steller sea lions by killer whales, the highest priority of the study is to estimate the abundance of transient whales. In the summer, Steller sea lions are thought to forage primarily in relatively close proximity to their rookeries and haulouts. Therefore, the study will be designed to include a 55-km area around Steller sea lion rookeries and major haulouts. The study area may need to be extended where rookeries and haulouts occur on small islands that are up to 20 km from the main coastline or the major islands. The eastern boundary of the study area is located at the eastern border of the western stock of Steller sea lions, along the coastline of the Kenai Peninsula (60°N, 150°W). The western boundary of the study area is located at the western side of Seguam Pass (56°N, 172°W) in the central Aleutian Islands. You will sample during three successive breeding seasons, 2008, 2009 and 2010, in order to improve the reliability of your estimates.

Problem 12: Impacts of Deer Browsing on Vegetation Regrowth

White-tailed deer (*Odocoileus virginianus*) subject many eastern forests in North America to high levels of herbivory. Deer typically browse seedlings and saplings of woody species during winter and predominantly herbaceous plants in the summer. Deer browsing can impact the growth and reproduction of both herbaceous and tree species and can also alter the structure of plant populations and the composition of plant communities, thus affecting the rate and direction of succession. In addition, changes in herbaceous and shrub layers due to deer browsing can indirectly influence tree species composition. Less palatable species, such as ferns and

grasses, often increase in abundance as a result of browsing and such changes can impact germination and establishment of tree seeds. Natural disturbances are common in most forests and a common consequence is a shift in plant species composition towards greater abundances of preferred browse species. As a result, deer prefer to feed in disturbed and early successional forest communities. Therefore, it is especially likely that deer herbivory may influence regeneration. Several researchers have studied the effects of deer browsing on logged areas or in old fields, but little is known about the potential impact of deer herbivory on natural forests as a result of browsing after natural disasters. You have an opportunity to fill this gap in our knowledge.

An F4 (wind speeds of over 430 km/h) tornado passed through a portion of Allegheny National Forest in Pennsylvania in 2000. The tornado caused complete canopy destruction along a 0.9 km wide swath of forest. The affected area lies within a 1,675-ha old growth preserve dominated by Eastern hemlock (*Tsuga Canadensis* (L.) Carr) and American beech (*Fagus grandifolia* Ehrh). Subdominant canopy species include sugar maple (*Acer saccharum* Marsh), red maple (*A. Rubrum* L.), yellow birch (*Betula alleghaniensis* Britton), and black cherry (*Prunus serotina* L.). The understory is predominantly *F. grandifolia* and *A. saccharum* saplings and the herbaceous layer includes *Oxalis acetosella* L., *Maianthemum canadense* Desf., *Arisaema triphyllum* (L.) Schott, and *Dryopteris spindosa* (Mueiler) Watt. Deer have been overabundant on the forested Allegheny Plateau in this part of Pennsylvania since the 1920s and by the mid-1930s, changes in abundance, height and species composition were observed for those plants within reach of deer. For many species, the abundance of seedlings greater than 0.3 m tall decreased significantly and *T. Canadensis* was virtually eliminated from the forest understories.

Problem 13: Emotional segregation in U.S. Films

The purpose of this study is to explore the persistence of a concept, known as "emotional segregation" through a systematic analysis of US films. Emotional segregation is defined as an institutionalized process, whereby European Americans are unable to see people of color as emotional equals or as capable of sharing the same human emotions and experiences. This concept was inspired by Charles C. Stember's (1976) study of sexual racism and Patricia Hill Collins' (2000) work on sexualized racism. In order to examine the existence of emotional segregation in the United States, a content analysis will be conducted of US films. Content analysis will be used to examine the extent of emotional segregation in US films made between 1980 and 2001 and containing interracial relationships between African Americans and 'whites', relationships between 'whites', and relationships between African Americans. These films will be selected from Leonard Maltin's 2001 *Movie and Video Guide*. This guide does not include every film made in the United States from 1980-2001. However, this guide is noted by *USA Today* as one of the most comprehensive works of its kind, consisting of over 20,000 film reviews, and is praised by the *New York Times* for including a wide range of films from blockbusters to 'little-known sleepers' and 'rarities'. Preliminary review of the films in the guide found that a total of 2,944 films included romantic relationships: 1,047 contained a relationship involving a 'white' man and woman, fifty-four contained a relationship involving an African American man and woman, seventeen contained a relationship involving an African American man and a 'white' woman, nineteen contained a relationship involving an African American woman and 'white' man, and 1,807 were 'other' films. Four main variables will be used to explore the main relationship in the film: success, duration, centrality, and the racial make-up of the relationship. Success and centrality are dichotomous variables that will be coded as either 'yes' or 'no' ('0' or '1'). The racial composition of the relationship (the independent variable) will include the following categories: 'black' male/'white' female (BM/WF), 'white' male/'black' female (WM/BF), 'white' male/'white' female (WW), and 'black' male/'black' female (BB). This variable is predicted

to directly affect the success, centrality, duration, and interactions of the relationship. Relationships will be defined as successful if they reach some degree of intimacy that was still present at the end of the film. Intimacy will be coded if the couple is married and/or when the characters confide personal thoughts to each other. In terms of measuring these variables, a coding sheet will be used that allows for the analysis of such interactions as embraces, kisses, touches, etc.

Problem 14: Taiwanese Consumers' Experiences Using Health-Related Websites

Electronic health commerce is playing a more and more significant role in the lives of individuals all over the world. According to one study, there were 2,293 websites providing health care information to the people of Taiwan in 2000. Internet use increased from about three million users in 1998 to approximately 8.6 million by the end of 2002. As in North America and Europe, more and more consumers rely on the Internet for health information in Taiwan. The purpose of your study is to explore Taiwan consumers' preferences and information needs, and the problems they encountered when getting information from medical websites. One of your research team will personally visit each **household** selected to participate in your study to complete an interview and ask the respondent to complete a short questionnaire. The visit will involve about 30 minutes with the respondent, as well as travel time to the individual's home. This can be very expensive and you have a limited budget. Therefore, devising a sampling approach that can provide the maximum number of completed interviews and questionnaires per dollar spent is very important to your study. You will conduct the study in Taipei. Taipei has 12 municipal districts. Figure 14 provides the name, total population, population density and area of each district in 2004.

Figure 14.1 Population, population density and area of the 12 districts of Taipei, 2004

District	Population	Population Density	Area (Sq. Km)
Beitou	248,965	4,382	56.82
Daan	312,956	27,546	11.36
Datong	127,022	22,357	5.68
Nangang	112,743	5,162	21.84
Neihu	263,651	8,349	31.58
Shilin	287,965	4,617	62.37
Songshan	208,832	20,870	9.29
Wanhua	194,876	21,958	8.85
Wenshan	260,412	8,264	31.51
Xinyi	231,015	20,612	11.21
Zhongshan	218,397	15,962	13.68
Zhongzheng	158,650	20,856	7.61