Assignment 3 Solutions

1. The question that the parents are asked is not relevant in determining the population. A random sample is always selected to be representative of the population. Since the sample consists of 100 parents with young children in Winnipeg, then the population is **all** parents with young children in Winnipeg.

2. Each digit on a random digit table has an equal chance of being chosen and is independent of any other digit that was previously selected. This means any set of four digits (including 0000 and 2847) has the same 1 in 10000 chance of being selected.

   No set of three digits is more likely to be observed than any other set of three digits, since digits appear on the table independently of any previous digits that have appeared.

   Although you may be surprised to see the five-digit number 77777, it is just as likely as any other five-digit number.

3. We'll read two digits at a time, ignoring the blanks. First, 29, which we skip because there no country labeled 29. Similarly, we skip 58. Then we will select 10 (Greece) for our first selection. Next, we skip 46, 37, and 51 and then select 19 (Poland) and 25 (Switzerland) for our second and third selections. Then we skip 60, 43, and 86. Next we come across 10, but 10 is already part of our sample, so we skip it this time. Then we skip 42 and select 27 (Ukraine) as the fourth selection. Skip 36 and then select 02 (Belgium) as the fifth selection. The next two digits are 19, which is already part of our sample, so we skip it. We also skip 53 and finally, 15 (Latvia) is our sixth selection.

4. Sampling scheme: "Random samples of 50 voters are contacted from each of the 13 provinces and territories."
   Type of sample: Stratified random sample. The country is subdivided into strata (the provinces and territories) and samples are taken from every stratum.

   Sampling scheme: "A random sample of five major Canadian cities is selected. Within each selected city, random samples of neighbourhoods are selected and all residents in the selected neighbourhoods are contacted."
   Type of sample: Multistage random sample. The first subdivision here is cities. Unlike what we would do with a stratified random sample, we don't select a sample from every city. Instead, we select a sample of cities. Then there are further levels of stratification after that.

   Sampling scheme: "From a list of all Canadian voters, 500 are randomly selected and contacted."
   Type of sample: Simple random sample. Here we are directly selecting from the population with a good sampling methodology.

   Sampling scheme: "Representatives from the polling firm select the first 50 voting-eligible adults they encounter at major public areas in each of the major Canadian cities."

Type of sample: Convenience sample. The sample is not selected by chance. Instead, individuals who are easiest to reach are chosen for the sample.

Sampling scheme: "In each of the major Canadian cities, the polling firm places an ad on the most popular radio station with a phone number to call to participate in the survey."
Type of sample: Voluntary response sample. The sample is not selected by chance. Respondents self-select themselves into the sample.

5. Situation: "Matthew had no chance of being included in the phone survey because he doesn't have a phone."
   Bias: Undercoverage. All individuals in the population should potentially be able to be selected as part of the sample. That is not the case here.

   Situation: "Natalie hung up the phone when she heard the person calling was from a polling firm."
   Bias: Nonresponse. In this case, Natalie may have been selected with an unbiased sampling methodology like simple random sampling, but she refuses to respond.

   Situation: "Jim phoned a number he saw on the 6:00 news to respond to a survey question."
   Bias: Voluntary response. Jim is self-selecting himself into the sample. The sample is not selected by chance.

6. We need to identify a variable that helps explain amount of jail time. "Seriousness of the crime" explains amount of jail time received.

7. If this were an experiment, the factor would be "type of club" and the factor levels (and treatments) would be "old clubs" and "new clubs". The experimental units would be each game of golf played (over both summers). However, an experiment requires that treatments are **randomly allocated** to the experimental units. That is not the case here. The first treatment "old clubs" are assigned to all of the experimental units (games) from the first summer and the second treatment "new clubs" are assigned to all of the games from the second summer. So the randomization principle of experimental design is not present here. This is simply an observational study.

8. Treatments are combinations of factor levels. There are 3 levels of the first factor, 2 levels of the second factor, and 2 levels of the third factor, so there are $3 \times 2 \times 2 = 12$ treatments.

9. The total number of subjects required is the number of treatments (12) times the number of subjects per treatment (5), so 60 subjects are required.

10. The only two factors here are amount of light and frequency of watering. The type of pine tree is not a factor – it is the blocking variable. Therefore, the factor levels are: full light, reduced light, watered once a week, no water.

11. Since we believe that responses may differ based on area of study, the blocking variable is area of study and there are three blocks, one for each area of study. Each block representing an area of study should consist of all subjects from that area of study.

12. This statement is true.

13. In a matched pairs design, we don't randomly pair up individuals. Instead, individuals who are similar in some way are paired up. Randomization is used in a matched pairs design to assign the treatments to each member of each pair.

14. The two individuals within a pair are similar to each other (they belong to the same block). Within each block, we have a completely randomized design (with just two subjects) and **all** of treatments need to be distributed to the individuals within the block (and this allocation is done randomly), so we can flip a coin and assign the first treatment to one individual and the second treatment to the other individual.