Set A 2) a] The two numbers for rate 28 vs 25 are very similar for Treatment a vaccine groups respectively. We are told that income and educational background are large factors in the rate of polio so the similarity of rate in the groups indicates that the families had similar income and educational background.

b] The numbers 4.6 vs 4.4 support this finding with the same reasoning as a)

c] The rates are very different 7.1 vs. 5.4, which (since no treatment was applied) indicates that the groups are different.

d] We know that individuals with lower socioeconomic backgrounds were both more likely to refuse treatment and less likely to get polio so the no-treatment group has lower socioeconomic status and lower rates of polio.

e] The groups are different. We already saw that there is a difference in socioeconomic status, and the groups may differ in other ways.
3) Individuals who knew they had the vaccine may be more likely to engage in common activities (like smoking) if they think they are protected by a vaccine.

7) The randomization did not work in ii) since the percent who smoked is drastically different between the control and treatment groups if it worked (like in i), the percentage should be very similar.

8) ii) Alcohol being a confounding variable means it is associated with both liver cancer and smoking.

12) We found that women who exercise more have fewer miscarriages. A potential confounding factor is age. Younger women are more likely to exercise more, and are less likely to have a miscarriage.

15) a) $\frac{46}{446} \cdot 100\% = 10.31\%$

b) Without more info it is reasonable to assume incomes are evenly distributed across this range $(10.31\%) \cdot (0.1) = 1.031\%$

c) 1.031%.

d) $1.031\% \cdot 2 = 2.062\%$
Chapter 3

Review 11.

a) Treatment - Prisoners who choose to attend the boot camp
   Control - Those who do not choose to attend

b) Observational Study - Prisons are not randomly assigned to treatment and control groups

c) False - We can not conclude this from an observational study. Potential confounding factors may include that prisoners who are motivated to stay out of jail may choose to attend the "boot camp."


<table>
<thead>
<tr>
<th></th>
<th>Ward 1</th>
<th>Ward 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Registered</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Democrat</td>
<td>1000</td>
<td>300</td>
</tr>
<tr>
<td>Republican</td>
<td>50</td>
<td>200</td>
</tr>
<tr>
<td>Vote</td>
<td>500</td>
<td>200</td>
</tr>
<tr>
<td>%</td>
<td>50%</td>
<td>100%</td>
</tr>
</tbody>
</table>

Total:

<table>
<thead>
<tr>
<th>Registered</th>
<th>Vote</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Democrat</td>
<td>1200</td>
<td>700</td>
</tr>
<tr>
<td>Republican</td>
<td>1200</td>
<td>750</td>
</tr>
</tbody>
</table>

The trend switches!
Exercise Set A

3) a) It is a probability method, we are using randomness to select our sample.

b) No, this is not the same as picking names out of a hat at random. Adjacent undergrads cannot be selected for example.

c) Probably not, but possibility to occur in rare situations. For example, a pair of jocks cannot be selected (right next to each other in alphabetical order).

3. ii) When we select a sample randomly, we tend to have a representative cross section of the population.