

Names of Group Members:

## 1 In Class Activity

### 1.1 Conditional Probability

Bob Ross was an American painter and a television host who is known for his landscape paintings with “*happy trees*” and encouraging quotes such as “*We don’t make mistakes, just happy little accidents*”. He hosted a television program called *The Joy of Painting* that aired from 1983 to 1994 on PBS. In each episode (except for two episodes) of this show he would complete one painting. The dataset we are going to use today has information on each of the paintings completed on this show. We are interested in two variables 1) whether a painting has a **tree** 2) whether the painting has a **mountain**. 1 indicates presence of the attribute and 0 indicates lack thereof.

```
bob_ross %>%  
  count(tree, mountain)
```

```
## # A tibble: 4 x 3  
##   tree mountain     n  
##   <int>     <int> <int>  
## 1     0         0   38  
## 2     0         1    4  
## 3     1         0  205  
## 4     1         1  156
```

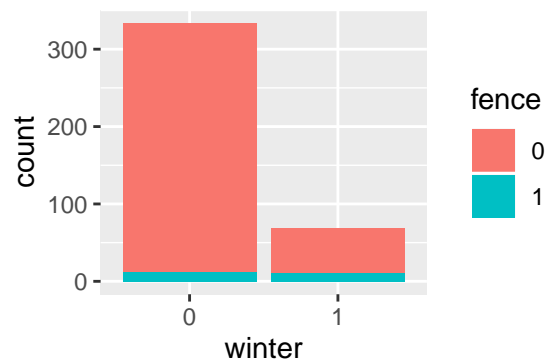
Q1. What is the probability that a randomly selected Bob Ross painting will have a tree in it?

Q2. What is the probability that a randomly selected Bob Ross painting will have a tree and a mountain?

Q3. You sat down and are watching an episode of *The Joy of Painting*. Bob Ross has already painted a mountain, what is the probability that this painting will have a tree?

Q4. Do the events of a painting having a tree and a painting having a mountain seem independent? Why or why not?

## 1.2 Visualizing Data



Q5. The figure above shows stacked bar plot for Bob Ross' paintings and whether these paintings have a winter scene and whether the painting has a fence. Fill in the blanks in the code below for the plot above.

```
bob_ross %>%  
  ggplot(aes(x = _____,  
             fill = _____)) +  
  _____
```

### 1.3 Discrete Distributions

x	1	2	3	4	5	6	$\geq 7$	<b>Total</b>
Number of re-spondents	346	910	431	399	169	47	46	2348
P(X=x) up to two dec- imal places								

Q6. The table above is taken from the General Social Survey that we have seen in previous examples in the class. Based on this data, calculate the expected household size as well as the variance.