

Homework Week 6

1. Using the data Irish pH data, generate a conditional dotplot, in which you will evaluate the relationship between pH values and the effect of the site being in a forested area or not.

(Hint: Create a categorical variable with the Forested variable with two levels (1,2) with labels (Forested, Non-forested), and create the dotchart - see the argument (groups) - Extracredit, the same dotchart in ggplot2 +1 point from this HW)

```

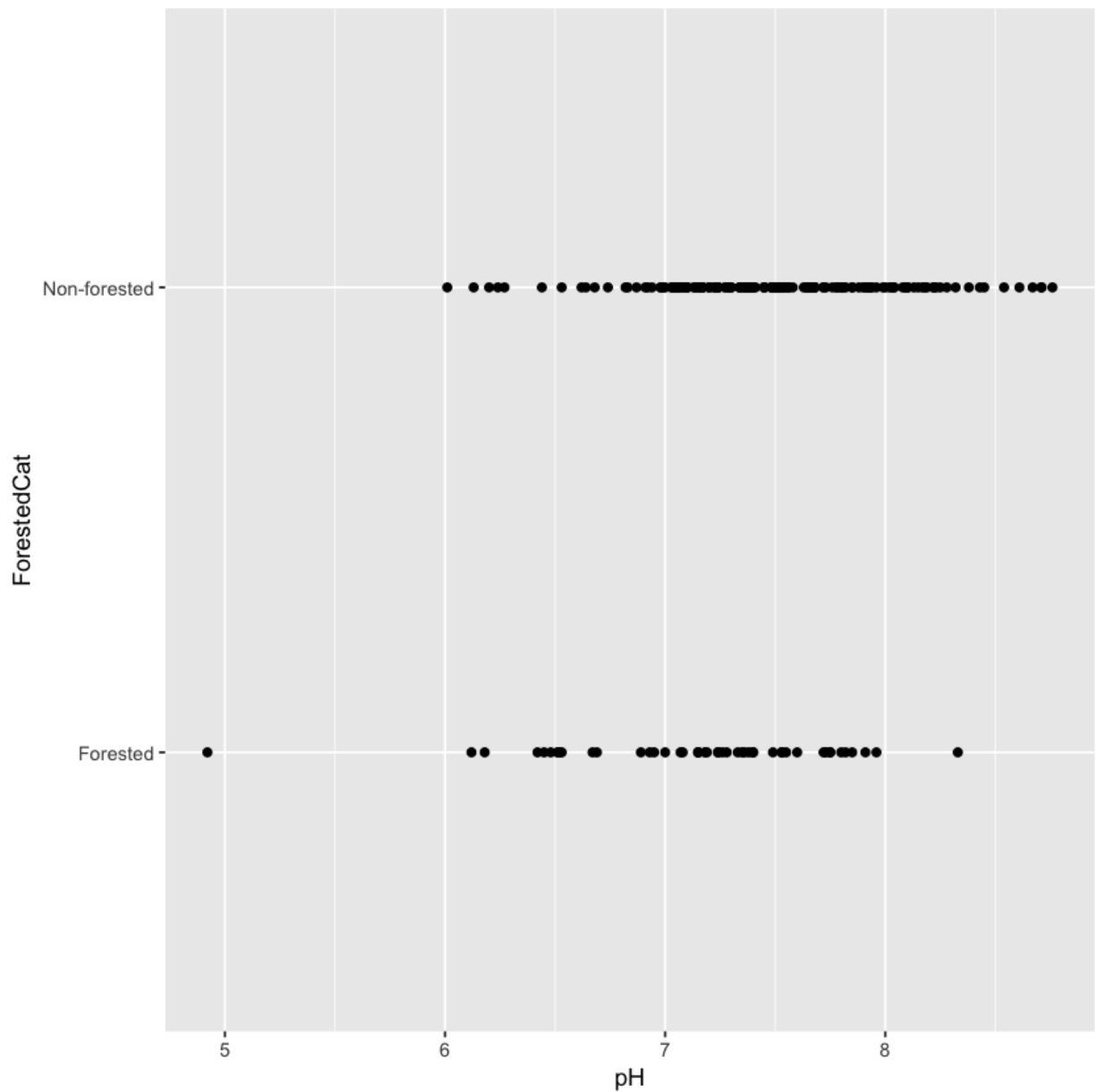
In [8]: library(ggplot2)
        IrishpH <- read.table(file = "IrishpH.txt",
                              header = TRUE,
                              dec = ".")

        IrishpH$ForestedCat[IrishpH$Forested == 1] = "Forested"
        IrishpH$ForestedCat[IrishpH$Forested == 2] = "Non-forested"

        ggplot(IrishpH, aes(ForestedCat,pH)) +
          geom_point(stat = "identity") +
          # geom_text(data=subset(IrishpH, Altitude > 400),
          #           aes(ID,Altitude,label=Altitude),hjust = -0.2)+
          coord_flip()

        #IrishpH

```



Non-forested area has a higher ph value. For example, the forested area only has one dot which ph value >8 , but for non-forested area, it has many dots which ph values >8 .

1. Find the following values: Φ represents the area under the curve (from the z table)

- a. $P(z < .79)$
- b. $P(z > .79)$
- c. $P(.39 < z < 1.39)$
- d. $P(.39 \leq z \leq 1.39)$
- e. $P(z < -0.79)$
- f. $P(z > -.79)$
- g. $P(-.79 < z < 1.39)$
- h. $P(-1.39 < z < -.79)$
- i. $P(-1.96 < z < 1.96)$
- j. $\Phi(1.96)$
- k. $\Phi(-1.96)$
- l. $\Phi(-1.08)$
- m. $1 - \Phi(1.08)$
- n. $\Phi(z) = .7157$

$$o. \Phi(z) = .3974$$

$$a. 0.7852$$

$$b. 1 - 0.7852 = 0.2148$$

$$c. P(z \leq 1.39) - P(z \leq -0.39) = 0.266$$

$$e. 0.2148$$

$$f. 1 - 0.2148 = 0.7852$$

$$g. P(z < -0.79) - P(z < -1.39) = 0.2148 - 0.0823 = 0.1325$$

$$i. P(z \leq 0.9750)$$

$$k. 0.0250$$

$$l. 0.1401$$

$$m. 1 - 0.8599 = 0.1401$$

$$n. z = 0.57$$

$$o. z = -0.26$$

1. Among females in the US between the ages of 18 and 74, diastolic blood pressure is normally distributed with mean 77 mm Hg and standard deviation of 11.6 mm Hg. What is the probability that a randomly selected woman has a diastolic blood pressure less than 60 mm Hg? greater than 90 mm Hg? between 60 and 90 mm Hg?

$$X \sim N(77, 134.56)$$

$$z_1 = (60 - 77) / 11.6 = -1.47$$

$$z_2 = (90 - 77) / 11.6 = 1.12$$

$$P(z < -1.47) = 0.0708$$

$$P(z > 1.12) = 1 - 0.8665 = 0.1335$$

$$P(-1.47 < z < 1.12) = P(z < 1.12) - P(z < -1.47) = 0.8665 - 0.0708 = 0.7957$$

1. The distribution of weights for the population of males in the United States is approximately normal with mean = 172.2 pounds and standard deviation of 29.8 lbs. What is the probability that a randomly selected man weights less than 130 lbs? More than 210 lbs? What is the probability that among five males selected at random from the population, at least one will have a weight outside the range 130 to 210 lbs?

$$X \sim N(172.2, 888.04)$$

$$z_1 = (130 - 172.2) / 29.8 = -1.42$$

$$z_2 = (210 - 172.2) / 29.8 = 1.27$$

$$P(z < -1.42) = 0.0778$$

$$P(z > 1.27) = 1 - 0.8980 = 0.102$$

$$p_1 = P(-1.42 < z < 1.27) = 0.8980 - 0.0778 = 0.8202 \quad P(\text{outside the range}) = 1 - 0.8202 = 0.1798 \quad P(\text{nobody outside the range}) = 0.8202^5 = 0.3712 \quad P(\text{at least one will have a weight outside the range}) = 1 - 0.3712 = 0.6288$$

