**R CODE (FULL SOLUTION TO BE POSTED AFTER 4PM)**

**QUESTION 2**

```
set.seed(246)
n <- 20
repetitions <- 1000
sim20 <- rep(NA, repetitions)
for (i in 1:repetitions)
{
new_sim <- sample(AutoClaims$AGE,size = n)
sim_mean <- mean(new_sim)
sim20[i] <- sim_mean
}
sim20 <- data_frame(means = sim20)
sim20 %>% ggplot(aes(x = means)) +
geom_histogram(binwidth = 2, colour = "black", fill = "grey") +
labs(x="Sample mean ages for samples of size 20")

summarise(sim20, min=min(means), mean = mean(means), median = median(means),
max=max(means), sd = sd(means), n=n())

set.seed(321)
ages20 <- data_frame(age=sample(AutoClaims$AGE,size = 20, replace=FALSE))
glimpse(ages20)

ages20 %>% ggplot(aes(x = age)) + geom_histogram(binwidth = 2, colour = "black", fill = "grey") +
labs(x="Ages of Claimants of Random Sample of 20 Claims")

summarise(ages20, min=min(age), mean = mean(age), median = median(age), max=max(age), sd =
sd(age), n=n())

set.seed(123)
boot_means <- rep(NA, 1000) # where we'll store the bootstrap means
sample_size <- 20
for (i in 1:1000)
{
boot_samp <- ages20 %>% sample_n(size = sample_size, replace=TRUE)
boot_means[i] <- as.numeric(boot_samp %>% summarize(mean(age)))
}

boot_means <- data_frame(means=boot_means)
boot_means %>% ggplot(aes(x = means)) +
geom_histogram(binwidth = 2, colour = "black", fill = "grey") +
labs(x="Sample mean ages for samples of size 20")

summarise(boot_means, min=min(means), mean = mean(means), median = median(means),
max=max(means), sd = sd(means), n=n())
```

**QUESTION 3**

```
Gestation <- Gestation %>% filter(!is.na(age))
summarize(Gestation, n=n())
boot_means <- rep(NA, 5000) # where we'll store the bootstrap means
sample_size <- as.numeric(Gestation %>% summarize(n()))
set.seed(50)
for (i in 1:5000)
{
boot_samp <- Gestation %>% sample_n(size = sample_size, replace=TRUE)
boot_means[i] <- as.numeric(boot_samp %>% summarize(mean(age)))
}
quantile(boot_means,c(.005,.995))

boot_medians <- rep(NA, 5000)
sample_size <- as.numeric(Gestation %>% summarize(n()))
set.seed(579)
for (i in 1:5000)
{
boot_samp <- Gestation %>% sample_n(size = sample_size, replace=TRUE)
boot_medians[i] <- as.numeric(boot_samp %>% summarize(median(age)))
}
quantile(boot_medians,c(0.025,0.975))
```