

A2-singhvis

Vishal Singh

September 11, 2019

```
library(tidyverse)
library(MASS)
library(dplyr)
```

1.

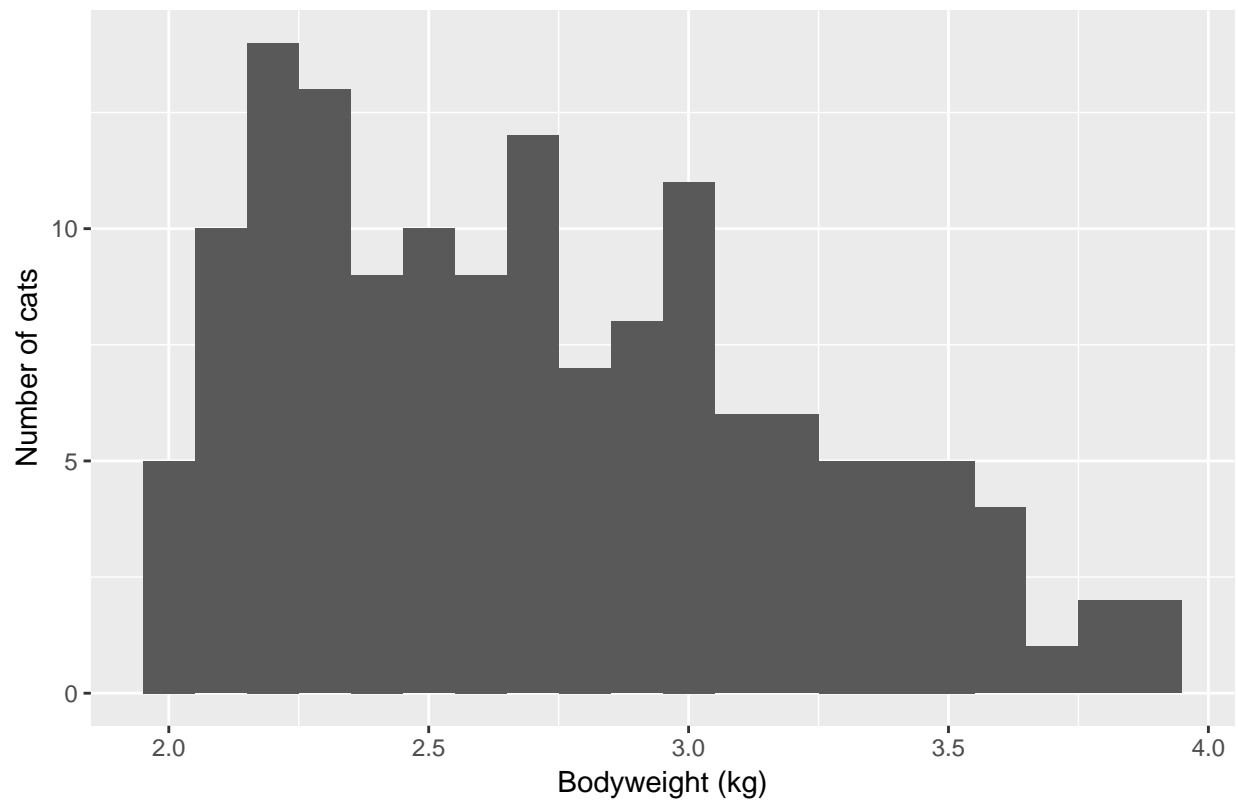
We can observe the following-

1. The distribution is right skewed
2. Both mean and median < 3 (~2.7kgs), i.e. most of the cats weigh less than 3 kgs
3. The range of weights of is from 2kgs - 3kgs
4. Most number of cats-14 weigh 2.3 kgs

```
cats.df = MASS::cats

ggplot(cats.df, aes(x = Bwt)) + geom_histogram(bins = 20) +
  xlab("Bodyweight (kg)") +
  ylab("Number of cats") +
  ggtitle("Cat weights from Fisher (1947)")
```

Cat weights from Fisher (1947)



2.

b.

From the plots we can observe the following-

1. The general feeling towards Hilary Clinton is that of “very cold” with maximum values near 0 (25th percentile being 3) and mean being 43. The feeling for Hilary Clinton is somewhat polarizing as most values are concentrated towards the extreme ends.
2. The plot for Donald Trump is similar to Hilary Clinton’s, will most values on extreme ends thus showing the views towards Donald Trumpb being polar. Mean for Donald Trump is the lowest (40) showing the sentiment towards Trump being coldest.
3. Most values for Marco Rubio are midway between very cold and very hot, there are a lot of values for “very cold” as well.
4. The plot for Bernie Sanders is the most distributed of all the candidates, there are comparable values for “very cold”, “very hot” and midway between the two
5. The mean values for Sanders (~57) is the highest which shows the general feeling is the warmest towards him whereas the mean values is lowest for Trump (~40) which shows the feeling is coldest towards Trump

```
anes_pilot_2016 = read.csv("C:\\Users\\singh\\Documents\\Acads\\EDA\\A2\\anes_pilot_2016.csv")

anes = anes_pilot_2016 %>% dplyr::select("fttrump", "fthrc", "ftsanders", "ftrubio", "immig_numb") %>%

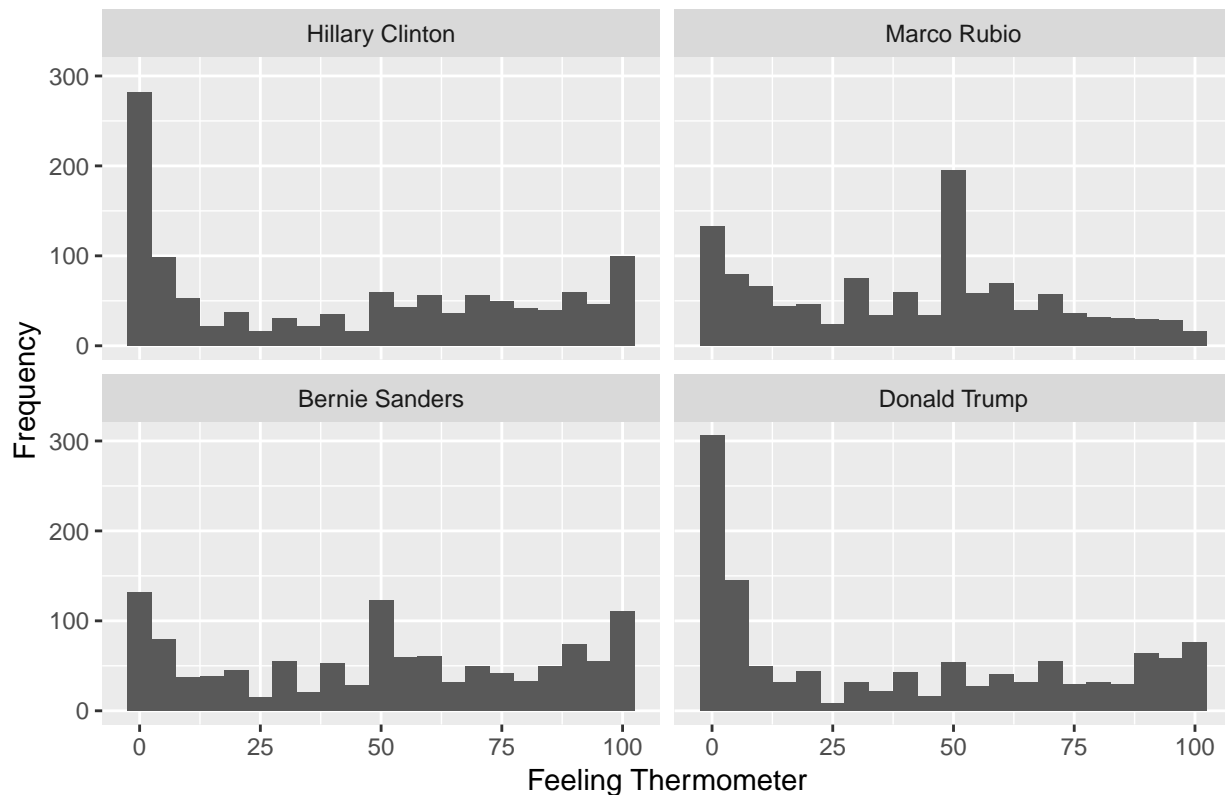
anes.long = gather(anes, key = "candidate", value="feeling_thermometer", c("fttrump", "fthrc", "ftsanders", "ftrubio"))

anes.long = anes.long[anes.long$feeling_thermometer <= 100,]

candidate.names = c("fttrump"="Donald Trump", "fthrc"="Hillary Clinton", "ftsanders"="Bernie Sanders", "ftrubio"="Marco Rubio")

ggplot(anes.long, aes(x = feeling_thermometer)) +
  geom_histogram(binwidth = 5) +
  facet_wrap(~ candidate, ncol = 2, labeller = as_labeller(candidate.names)) +
  xlab("Feeling Thermometer") +
  ylab("Frequency") +
  ggtitle("Distribution of Feeling Thermometer for Different Candidates")
```

Distribution of Feeling Thermometer for Different Candidates



b.

We can infer the following from the plots of-

1. The mean feeling for people who feel immigration should be increased a lot is highest for Hillary Clinton(60) and Bernie Sanders(64)(Democrats), this means that people who feeling immigration should be incresed lot have a warmer feeling towards Bernie and Hilary. There is a stark difference between means of Hilary/Bernie(democrats) and Trump/Rubio(republicans) (~60 vs 20-30)
2. The trend can be seen for categories who feeling immigration should be increased in some amount. The means for democrats is higher than republicans in these categories
3. The mean for categories which feeling immigration should be decreased (in varying amounts) is higher for republicans, this means that people who feel immigration should be decreased have a warmer feeling towards republicans
4. The difference in means for republicans vs democrats for category “decreased a lot” is the highest. With mean of republican being 42, 60 vs 26, 30 being means for democrats

```
anes.long.mean = anes.long %>%
  group_by(immig_num, candidate) %>%
  summarise(Mean = mean(feeling_thermometer))

ggplot(anes.long.mean, aes(x = candidate, y = Mean)) +
  geom_bar(stat = 'identity') +
  scale_x_discrete(labels = candidate.names) +
  facet_wrap(~ immig_num, nrow = 3) +
  coord_flip() +
  xlab("Mean Feeling Thermometer values") +
  ylab("Immigration Responses") +
```

```
ggtitle("Mean Feeling Thermometer values by Immigration  
Responses for different Candidates")
```

