Figure 1
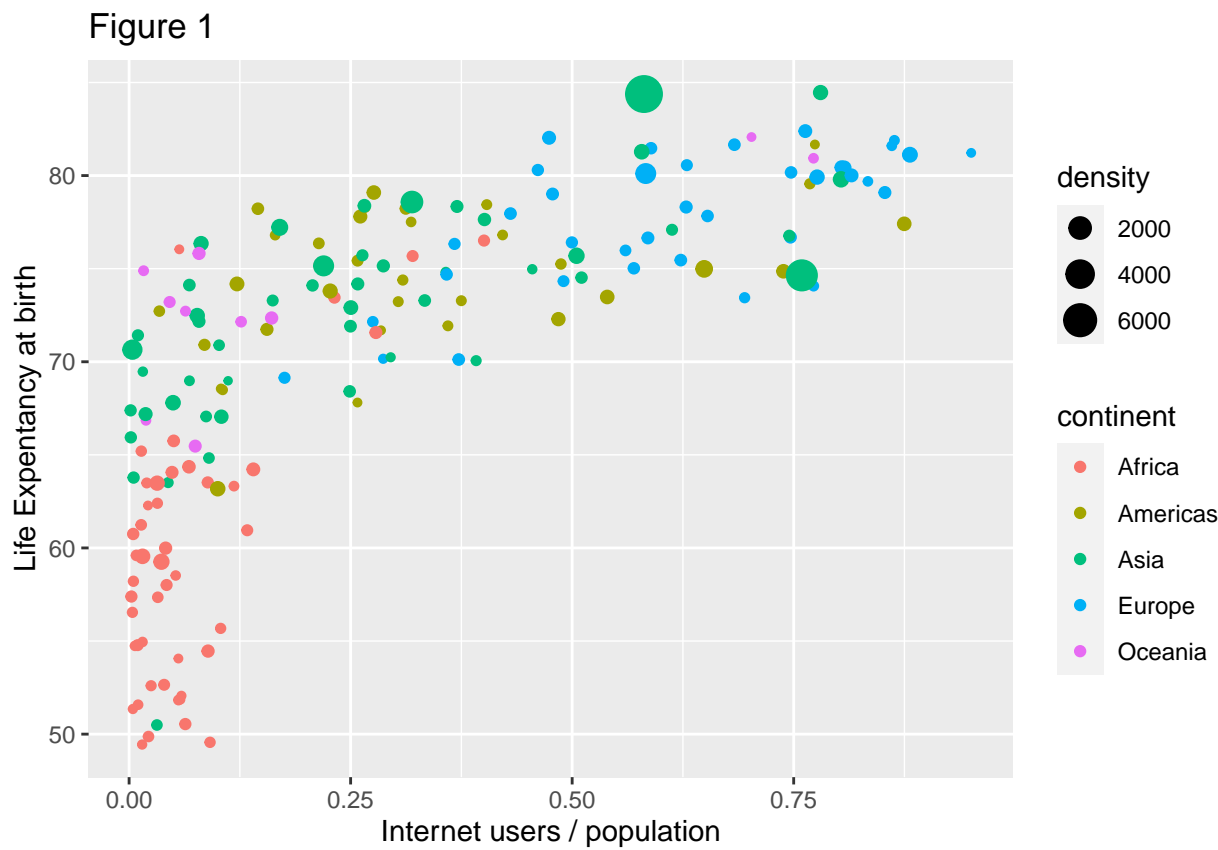
Figure 1

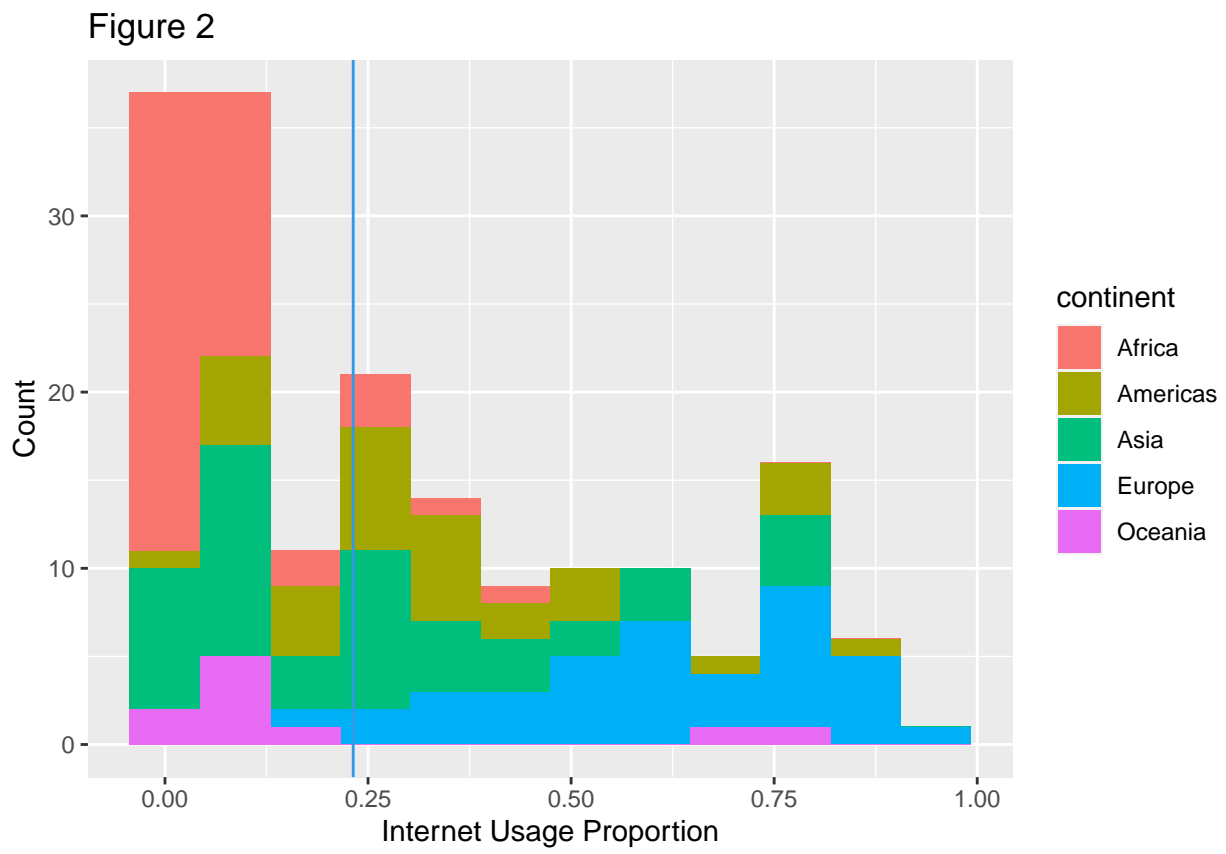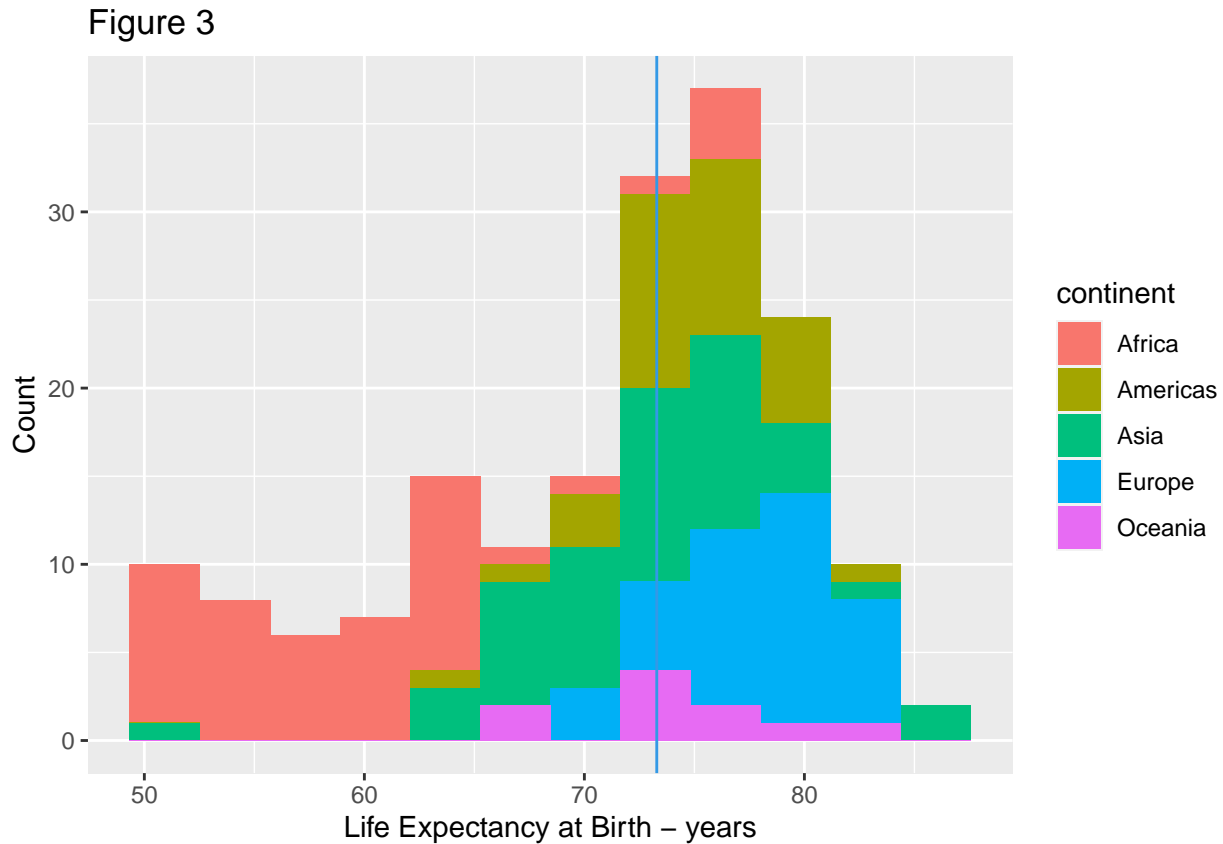**Figure 2**

Figure 2

Figure 3
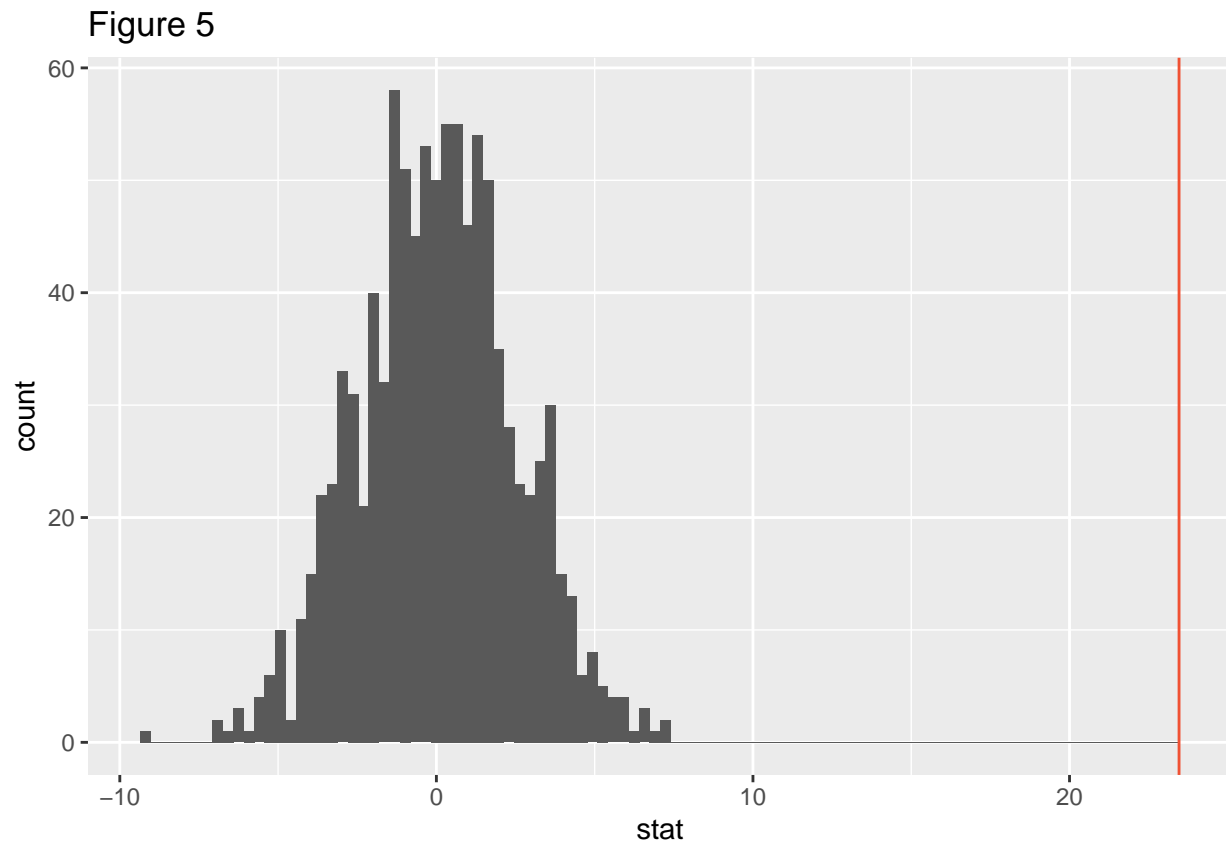

Figure 3

Figure 4

```
lm_mod <- lm(life_exp_at_birth ~ internet_usage_proportion, data = cia)
obs_slope <- lm_mod$coefficients[2]
summary(lm_mod)
```

```
##
## Call:
## lm(formula = life_exp_at_birth ~ internet_usage_proportion, data = cia)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -16.2903  -2.8736   0.1377   4.4186  11.1027
##
## Coefficients:
##                           Estimate Std. Error t value Pr(>|t|)
## (Intercept)                63.7069     0.6732   94.63   <2e-16 ***
## internet_usage_proportion  23.4594     1.6731   14.02   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 6.066 on 175 degrees of freedom
## Multiple R-squared:  0.5291, Adjusted R-squared:  0.5264
## F-statistic: 196.6 on 1 and 175 DF,  p-value: < 2.2e-16
```

Figure 5

```
set.seed(1009)
perm_slope <- cia %>%
    specify(life_exp_at_birth ~ internet_usage_proportion) %>%
    hypothesize(null = "independence") %>%
    generate(reps = 1000, type = "permute") %>%
    calculate(stat = "slope")

ggplot(data=perm_slope, aes(x=stat)) +
    geom_histogram(bins=100) +
    geom_vline(xintercept = obs_slope,
              color = "#F05133") + ggtitle("Figure 5")
```


Figure 5

```
    xlim(-30,30)
```

```
## <ScaleContinuousPosition>
##  Range:
##  Limits:  -30 --   30
```

Figure 6

```
lm_res <- resid(lm_mod)
lm_res <- data.frame(resid = lm_res)
lm_res <- lm_res %>% mutate(internet_usage_proportion = cia$internet_usage_proportion)
histogram <- ggplot(data=lm_res, aes(x=resid))
histogram + geom_histogram(binwidth=1, color="black") +
```

```
xlab("Residuals") +  ylab("Frequency") + ggtitle("Figure 6: Histogram of Residuals")
```
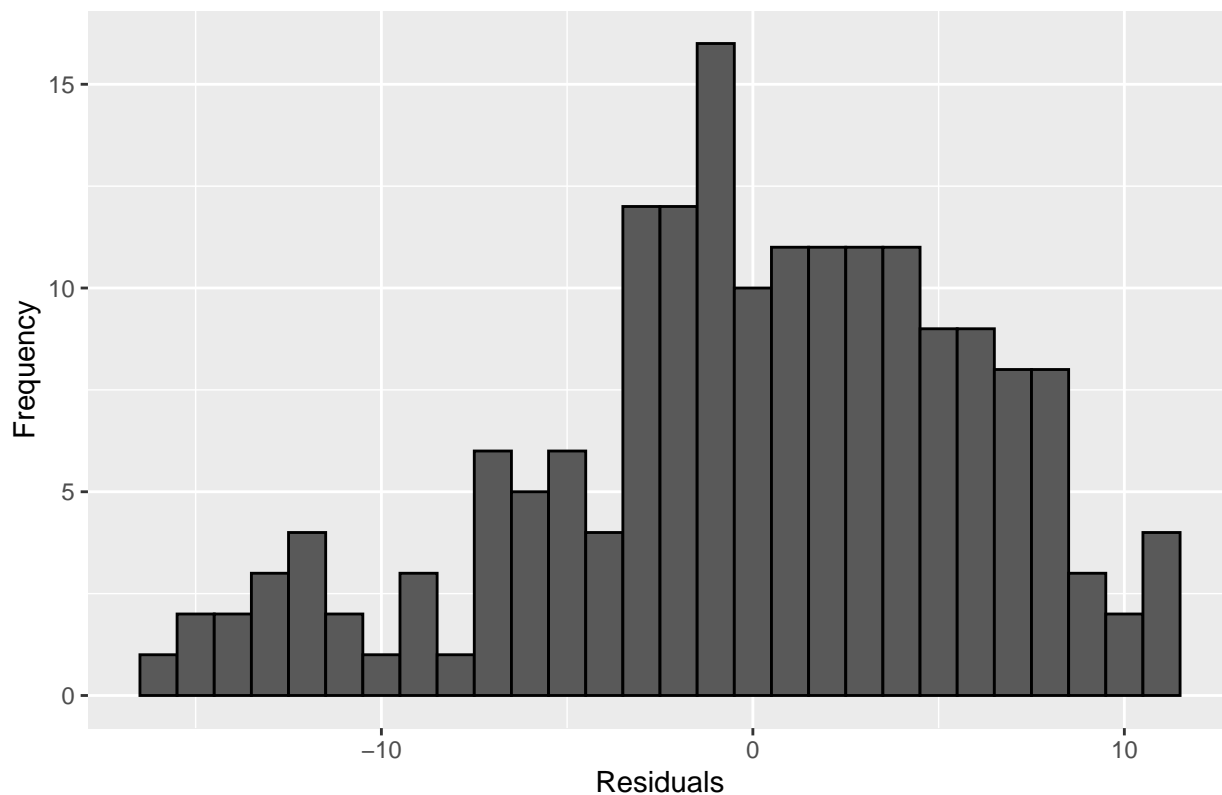
Figure 6: Histogram of Residuals

Figure 7

```
scatter <- ggplot(data=lm_res, aes(x = internet_usage_proportion, y =resid))
scatter + geom_point(size = 1.5) +
  xlab("Internet Usage Proportion") +  ylab("Residual")  +
  geom_hline(yintercept=0, linetype='dashed', col = 'red') +
  ggtitle("Figure 7: Internet Usage Proportion vs Residuals")
```
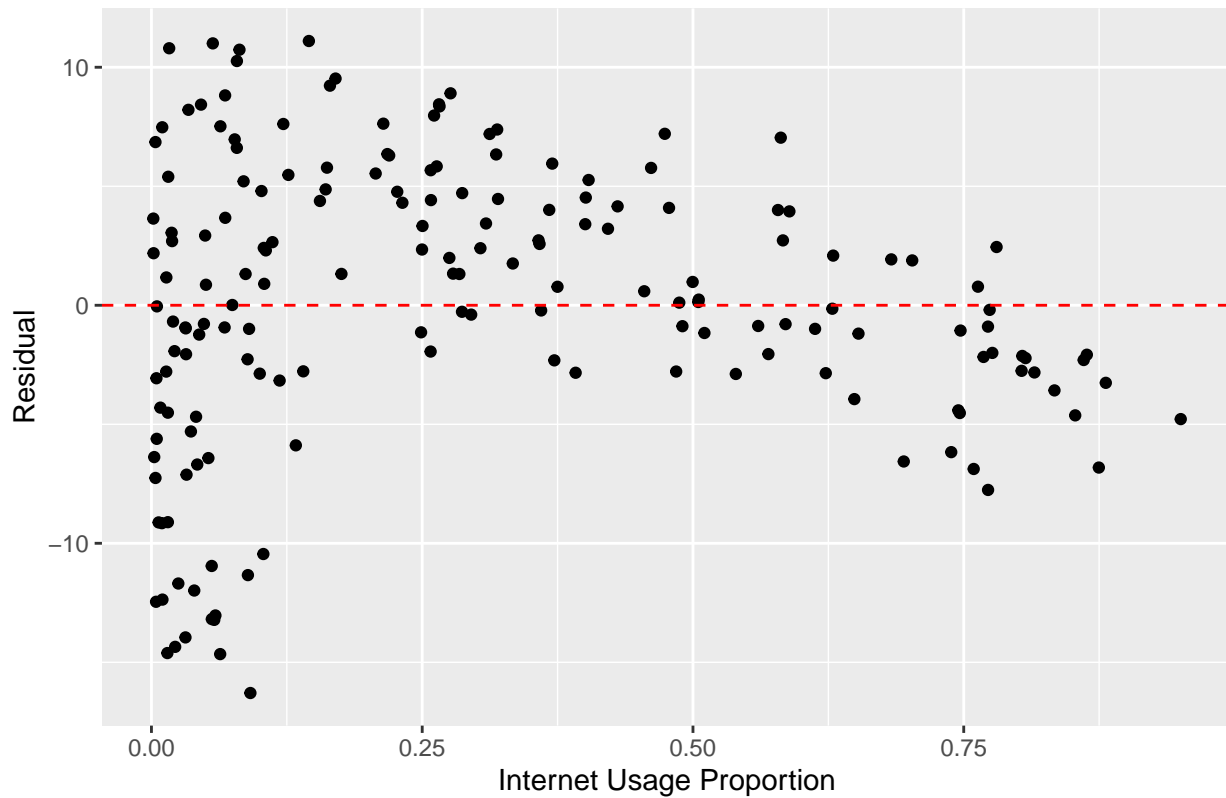
Figure 7: Internet Usage Proportion vs Residuals

Figure 8

```
set.seed(34209)
perm_ci <- cia %>%
  specify(life_exp_at_birth ~ internet_usage_proportion) %>%
  generate(reps = 1000, type = "bootstrap") %>%
  calculate(stat = "slope")
visualize(perm_ci)
```
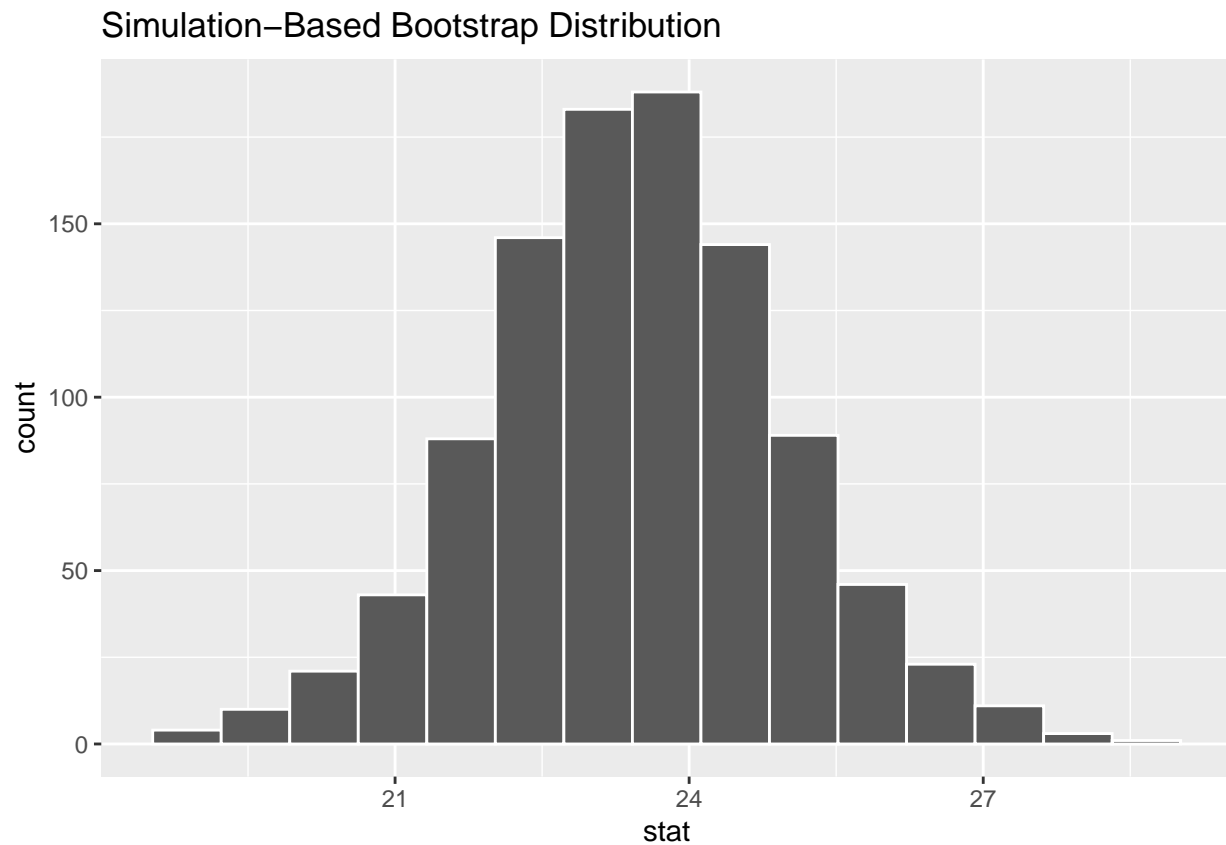
Simulation–Based Bootstrap Distribution

Figure 9

```
alpha <- .05

#lower percentile cutoff
p_lower <- .025

#upper percentile cutoff
p_upper <- 1-(.025)

# Create a confidence interval of stat using quantiles
quantile(perm_ci$stat,c(p_lower,p_upper))

##     2.5%    97.5%
## 20.39227 26.50567
```

Figure 10

```
noafrica <- cia %>%
  filter(continent != "Africa")

ggplot (data=noafrica, aes(x=internet_usage_proportion, y=life_exp_at_birth)) +
  geom_point(aes(color = continent,size=density)) +
  xlab('Internet users / population') +
  ylab('Life Expentancy at birth') +
  ggtitle('Figure 10:  Africa Excluded')
```
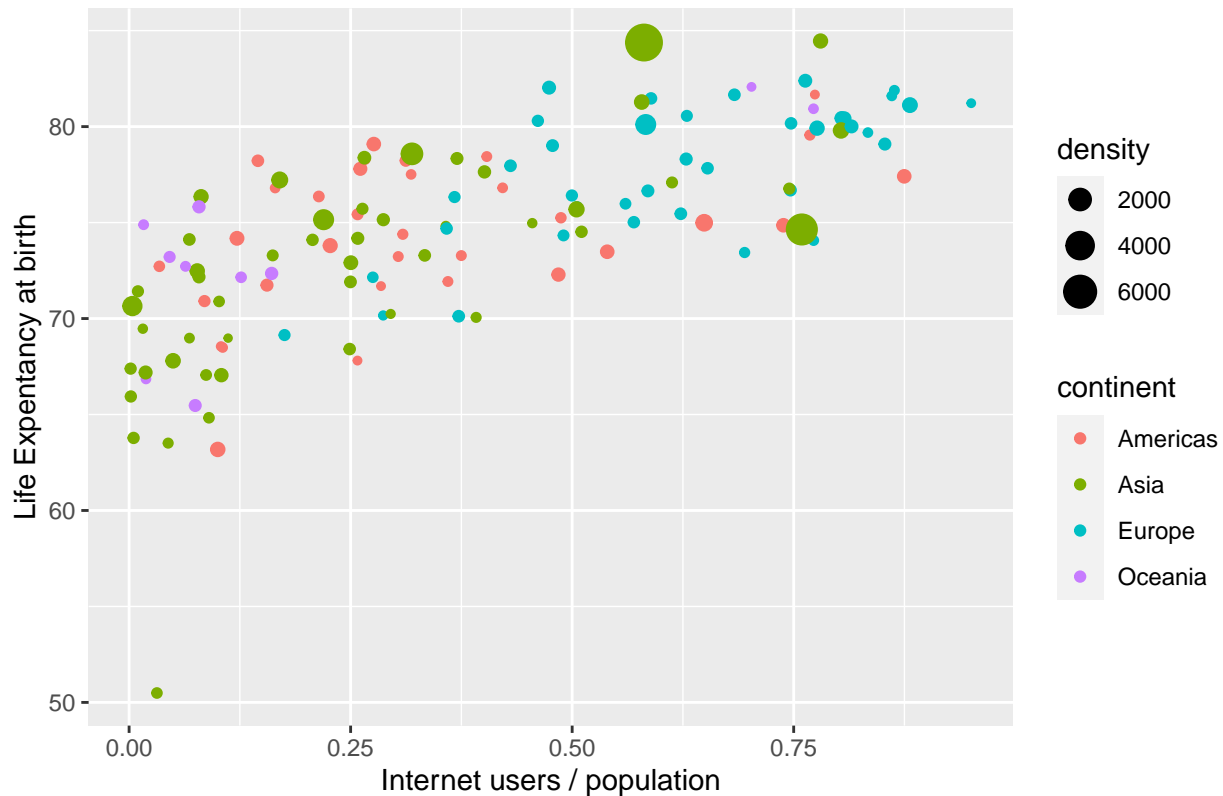
Figure 10: Africa Excluded

Figure 11

```
lm_mod_noafrica <- lm(life_exp_at_birth ~ internet_usage_proportion, data=noafrica)
obs_slope_noafrica <- lm_mod_noafrica$coefficients[2]
obs_slope_noafrica
```

```
## internet_usage_proportion
##                  13.52906
```

```
summary(lm_mod_noafrica)
```

```
##
## Call:
## lm(formula = life_exp_at_birth ~ internet_usage_proportion, data = noafrica)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -19.4508  -2.3693   0.0082   2.6411   7.0030
##
## Coefficients:
##                            Estimate Std. Error t value Pr(>|t|)
## (Intercept)                 69.5157     0.5602  124.08   <2e-16 ***
## internet_usage_proportion   13.5291     1.2007   11.27   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 3.678 on 127 degrees of freedom
```

```
## Multiple R-squared:  0.4999, Adjusted R-squared:  0.496
## F-statistic:   127 on 1 and 127 DF,  p-value: < 2.2e-16
```

Figure 12

```
set.seed(1010)
perm_slope_noafrica <- noafrica %>%
   specify(life_exp_at_birth ~ internet_usage_proportion) %>%
   hypothesize(null = "independence") %>%
   generate(reps = 1000, type = "permute") %>%
   calculate(stat = "slope")

ggplot(data=perm_slope_noafrica, aes(x=stat)) +
   geom_histogram(bins=100) +
   geom_vline(xintercept = obs_slope_noafrica,
              color = "#F05133") +
   ggtitle("Figure 12")
```
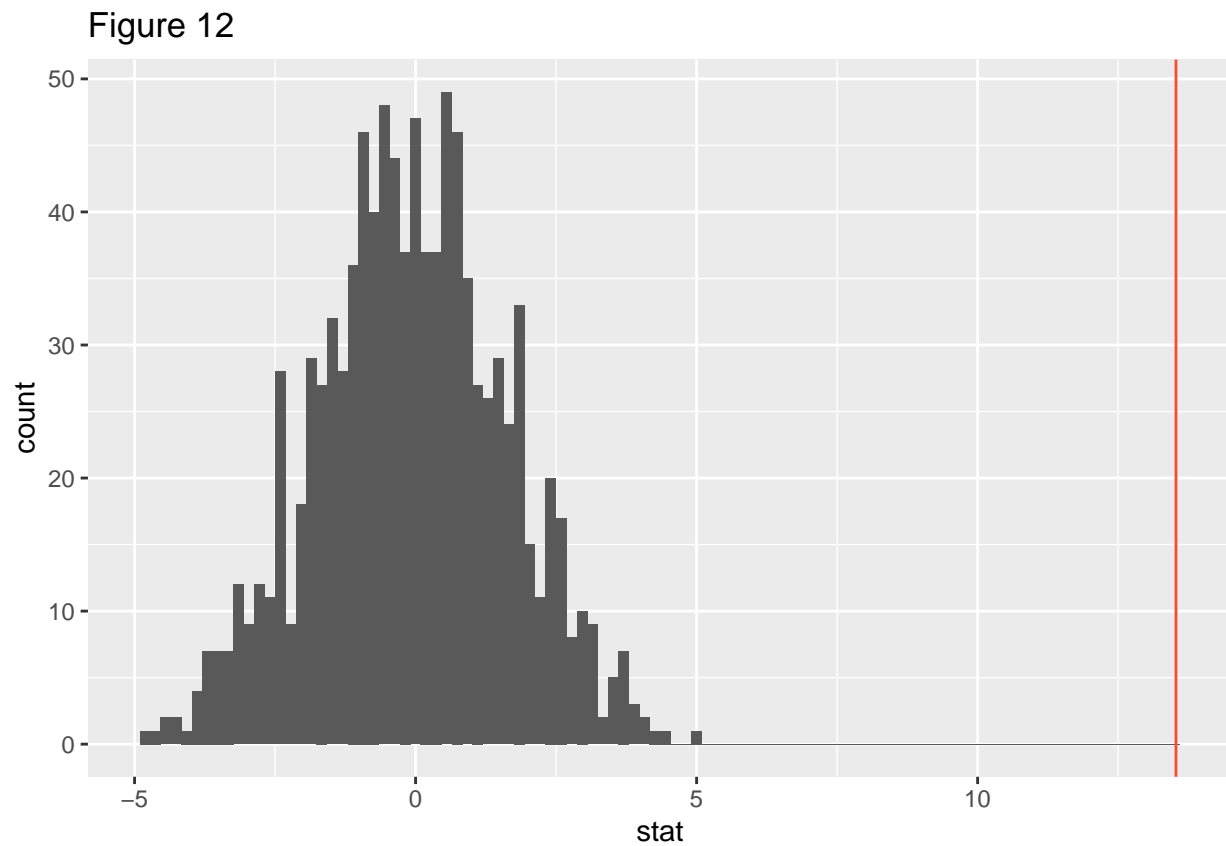


```
   xlim(-20,20)
```

```
## <ScaleContinuousPosition>
##  Range:
##  Limits:  -20 --   20
```

Figure 13

```
set.seed(34210)
perm_ci_noafrica <- noafrica %>%
  specify(life_exp_at_birth ~ internet_usage_proportion) %>%
  generate(reps = 1000, type = "bootstrap") %>%
  calculate(stat = "slope")
visualize(perm_ci_noafrica)
```
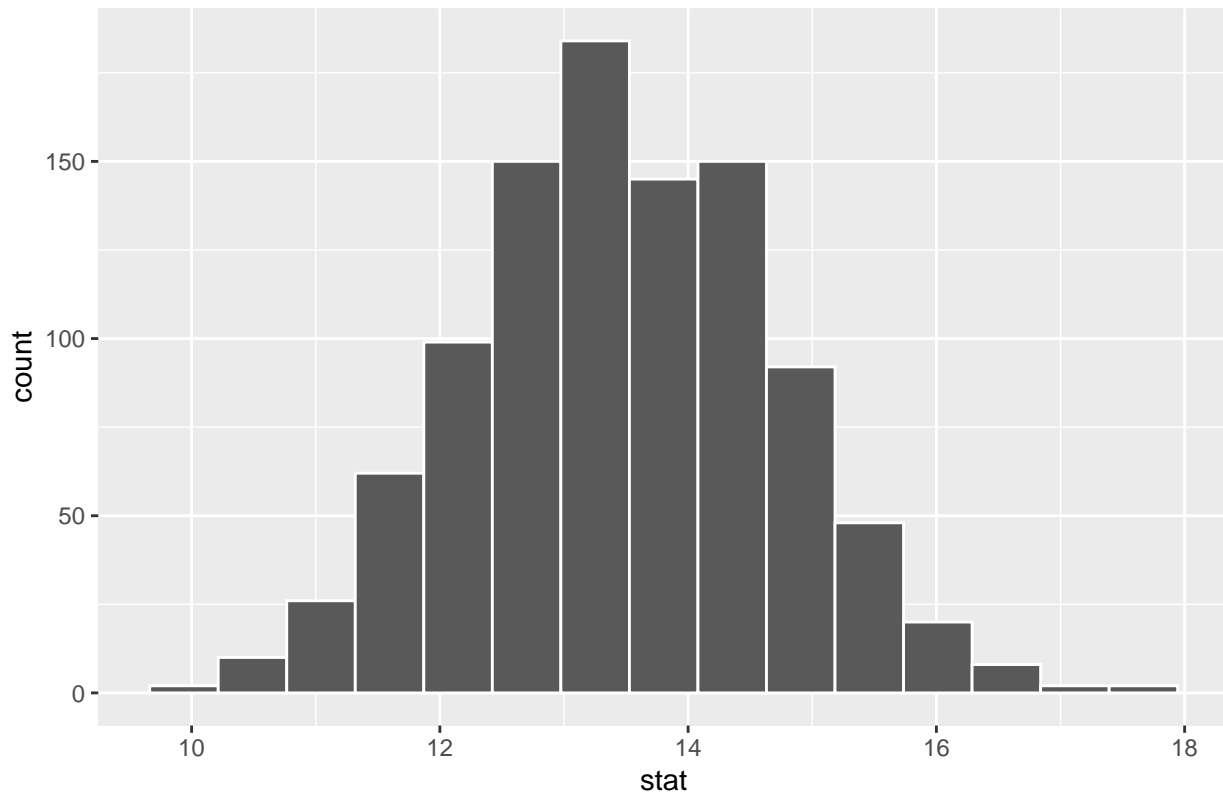

Figure 14

```
alpha <- .05

#lower percentile cutoff
p_lower <- .025

#upper percentile cutoff
p_upper <- 1-(.025)

# Create a confidence interval of stat using quantiles
quantile(perm_ci_noafrica$stat,c(p_lower,p_upper))

##     2.5%    97.5%
## 11.10236 15.85847
```

Figure 15

```
lm_res_noafrica <- resid(lm_mod_noafrica)
lm_res_noafrica <- data.frame(resid = lm_res_noafrica)
lm_res_noafrica <- lm_res_noafrica %>% mutate(internet_usage_proportion = noafrica$internet_usage_propo
histogram <- ggplot(data=lm_res_noafrica, aes(x=resid))
histogram + geom_histogram(binwidth=1, color="black") +
  xlab("Residuals") +  ylab("Frequency") + ggtitle("Figure 15: Histogram of Residuals with Africa Exclu
```


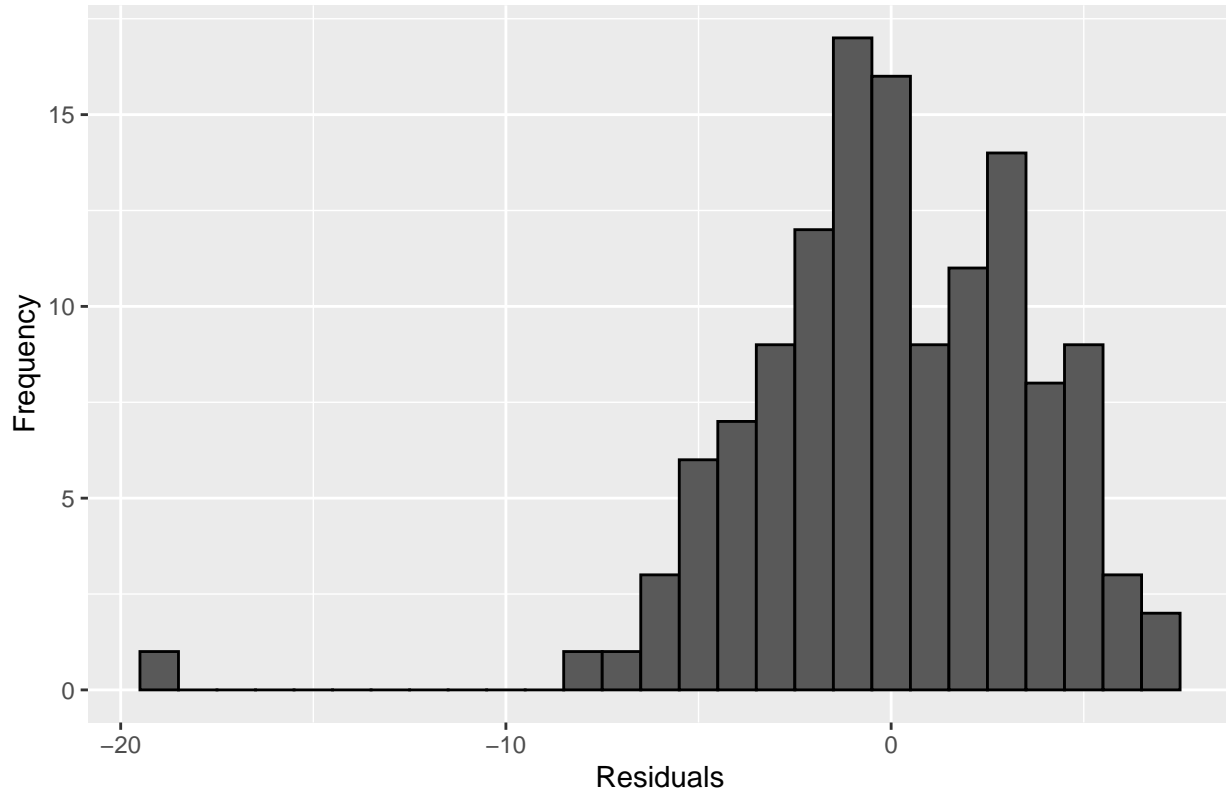Figure 15: Histogram of Residuals with Africa Excluded

Figure 16

```
scatter <- ggplot(data=lm_res_noafrica, aes(x = internet_usage_proportion, y =resid))
scatter + geom_point(size = 1.5) +
  xlab("Internet Usage Proportion") +  ylab("Residual")  +
  geom_hline(yintercept=0, linetype='dashed', col = 'red') +
  ggtitle("Figure 16: Internet Usage Proportion vs Residuals with Africa Excluded")
```

Figure 16: Internet Usage Proportion vs Residuals with Africa Excluded