Renewvia - Survey Impact Analysis

Importing the libraries and datsets

```
tinytex::install_tinytex()
library(dataMaid)
library(ggplot2)
library(dplyr)
```

```
##
## Attaching package: 'dplyr'

## The following object is masked from 'package:dataMaid':
##
##     summarize

## The following objects are masked from 'package:stats':
##
##     filter, lag

## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

```
library(lessR)
```

```
##
## lessR 4.2.9                         feedback: gerbing@pdx.edu
## ------------------------------------------------------------
## > d <- Read("")   Read text, Excel, SPSS, SAS, or R data file
##   d is default data frame, data= in analysis routines optional
##
## Learn about reading, writing, and manipulating data, graphics,
## testing means and proportions, regression, factor analysis,
## customization, and descriptive statistics from pivot tables
##   Enter:  browseVignettes("lessR")
##
## View changes in this and recent versions of lessR
##   Enter: news(package="lessR")
##
## Interactive data analysis
##   Enter: interact()
```

```
##
## Attaching package: 'lessR'
```

```
## The following objects are masked from 'package:dplyr':
##
##     recode, rename
```

```
library(tidyr)
library(ggpubr)
library(exact2x2)
```

```
## Loading required package: exactci
```

```
## Loading required package: ssanv
```

```
## Loading required package: testthat
```

```
##
## Attaching package: 'testthat'
```

```
## The following object is masked from 'package:tidyr':
##
##     matches
```

```
## The following object is masked from 'package:dplyr':
##
##     matches
```

```
# Loading the paired data
initial_clean <- read.csv("data_cleaning/datasets_clean/initial_clean.csv")
initial_enc <- read.csv("data_cleaning/datasets_encoded/initial_encoded.csv")
# makeDataReport(initial_clean,
#                output = "html",
#                replace = TRUE)
# create_report(df)
# ExpReport(df, op_file = 'SmartEDA_df.html')
```

```
hs_post_clean <- read.csv("data_cleaning/datasets_clean/hs_post_clean.csv")
hs_post_enc <- read.csv("data_cleaning/datasets_encoded/hs_post_encoded.csv")
# makeDataReport(hs_post_clean,
#                output = "html",
#                replace = TRUE)
```

```
ci_post_clean <- read.csv("data_cleaning/datasets_clean/ci_post_clean.csv")
ci_post_enc <- read.csv("data_cleaning/datasets_encoded/ci_post_encoded.csv")
# makeDataReport(ci_post_clean,
#                output = "html",
#                replace = TRUE)
```

```
initial_clean_id <- initial_clean %>%
                    filter(renewvia_id != "") %>%
                    filter(renewvia_id != "no_meter")
hs_post_clean_id <- hs_post_clean %>%
                    filter(renewvia_id != "") %>%
                    filter(renewvia_id != "no_meter")
ci_post_clean_id <- ci_post_clean %>%
                    filter(renewvia_id != "") %>%
                    filter(renewvia_id != "no_meter")
clean_paired <- merge(initial_clean_id, hs_post_clean_id,
                      suffix = c("_pre", "_post"),
                      by= 'renewvia_id')
```

```
initial_enc_id <- initial_enc %>%
                    filter(renewvia_id != "") %>%
                    filter(renewvia_id != "no_meter")
hs_post_enc_id <- hs_post_enc %>%
                    filter(renewvia_id != "") %>%
                    filter(renewvia_id != "no_meter")
ci_post_enc_id <- ci_post_enc %>%
                    filter(renewvia_id != "") %>%
                    filter(renewvia_id != "no_meter")
enc_paired <- merge(initial_enc_id, hs_post_enc_id,
                    suffix = c("_pre", "_post"),
                    by= 'renewvia_id')
```

1- GENDER EQUALITY

Data set: Household Post-Connection Survey

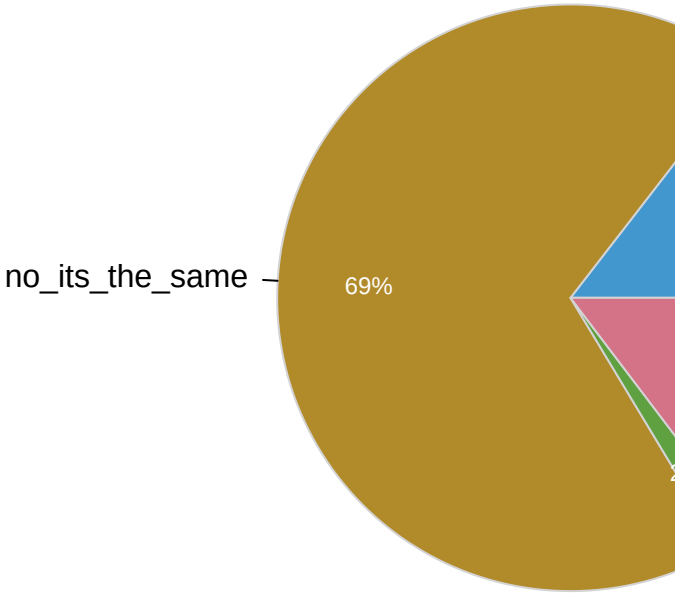
Statistical Test: Pearson's Chi-Square Goodness of fit

Hypothesis Testing:

- H_0: No changes in girls' school enrollment; H_1: A change in girls' school enrollment
- H_0: No changes in boys' school enrollment; H_1: A change in boys' school enrollment

```
# Breakdown of observed change
PieChart(girls_schooling_change, hole = 0, values = "%",
         data = hs_post_clean,
         main = "")
```

Question 1-a: Has there been a change in the number of females in your household who attend



school full time since connection to mini-grid power?

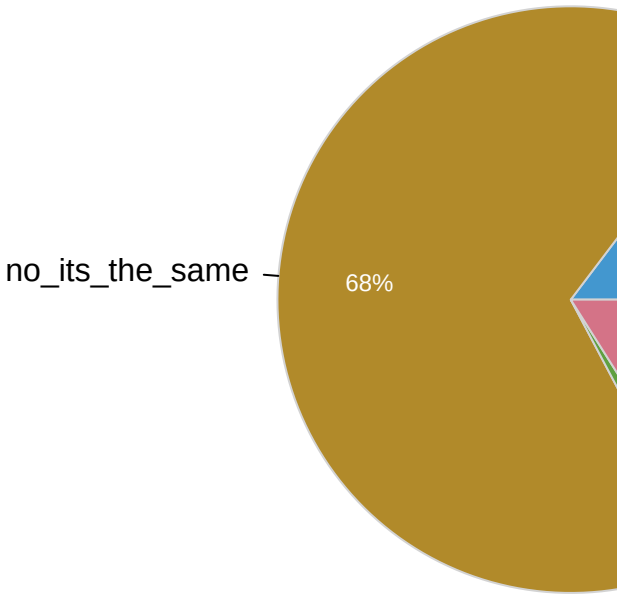
```
## >>> suggestions
## PieChart(girls_schooling_change, hole=0)  # traditional pie chart
## PieChart(girls_schooling_change, values="%")  # display %'s on the chart
## PieChart(girls_schooling_change)  # bar chart
## Plot(girls_schooling_change)  # bubble plot
## Plot(girls_schooling_change, values="count")  # lollipop plot
##
## --- girls_schooling_change ---
##
##     grls_schlng_c  Count   Prop
## --------------------------------
##                       378   0.145
##   no_its_the_same    1799   0.691
## yes_its_decreased      44   0.017
## yes_its_increased     382   0.147
## --------------------------------
##             Total    2603   1.000
##
## Chi-squared test of null hypothesis of equal probabilities
##   Chisq = 2817.123, df = 3, p-value = 0.000
```

```
girls_enrolled <- factor(hs_post_enc$girls_schooling_change)
girls_enrolled_table <-table(girls_enrolled)
chisq.test(girls_enrolled_table,p=rep(1/3,3))
```

```
##
##  Chi-squared test for given probabilities
##
## data:  girls_enrolled_table
## X-squared = 2338, df = 2, p-value < 0.00000000000000022
```

```
# Breakdown of observed change
PieChart(boys_schooling_change, hole = 0, values = "%",
         data = hs_post_clean_id,
         main = "")
```

Question 1-b: Has there been a change in the number of males in your household who attend



school FULL TIME since connection to mini-grid power?

```
## >>> suggestions
## PieChart(boys_schooling_change, hole=0)  # traditional pie chart
## PieChart(boys_schooling_change, values="%")  # display %'s on the chart
## PieChart(boys_schooling_change)  # bar chart
## Plot(boys_schooling_change)  # bubble plot
## Plot(boys_schooling_change, values="count")  # lollipop plot
##
## --- boys_schooling_change ---
##
##     bys_schlng_ch  Count   Prop
```

```
## --------------------------------
##                       382   0.147
##   no_its_the_same    1771   0.681
## yes_its_decreased      32   0.012
## yes_its_increased     417   0.160
## --------------------------------
##             Total    2602   1.000
##
## Chi-squared test of null hypothesis of equal probabilities
##   Chisq = 2712.801, df = 3, p-value = 0.000
```

```
boys_enrolled <- factor(hs_post_enc$boys_schooling_change)
boys_enrolled_table <-table(boys_enrolled)
chisq.test(boys_enrolled_table,p=rep(1/3,3))
```

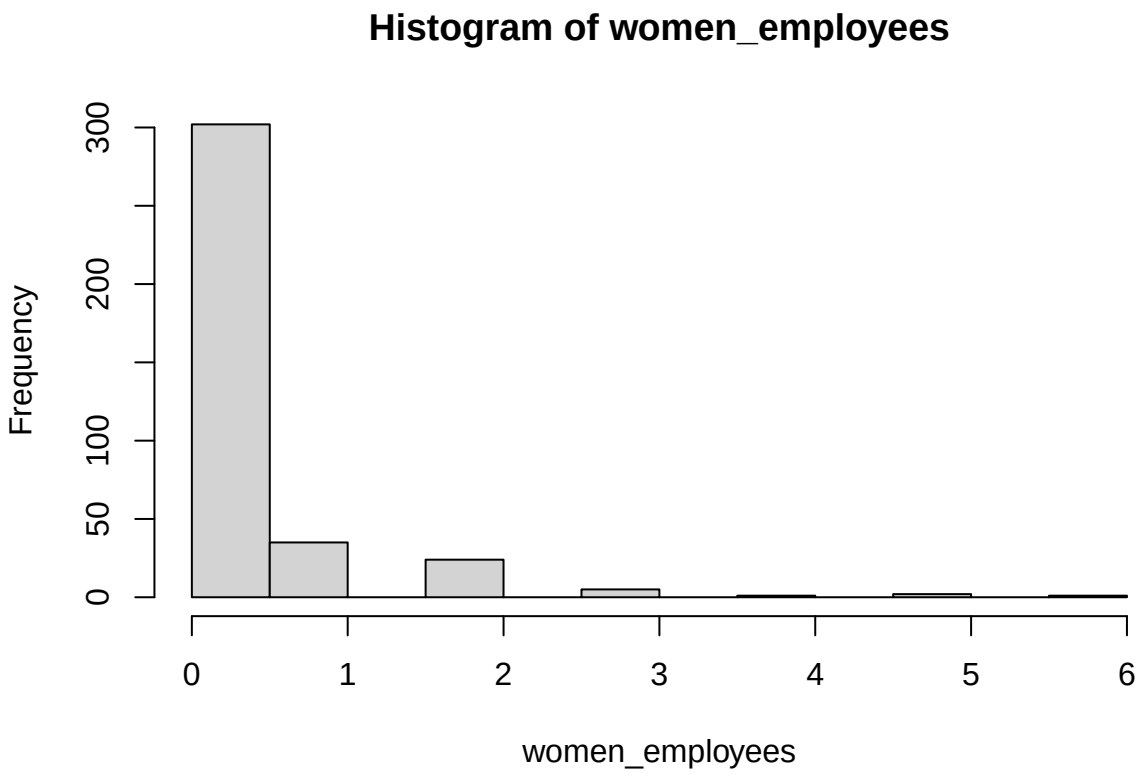
```
##
##  Chi-squared test for given probabilities
##
## data:  boys_enrolled_table
## X-squared = 2254.8, df = 2, p-value < 0.00000000000000022
```

Question 1-c: If you answered yes to adding new workers, how many new employees are female? Data set: Commercial Post-Connection Survey

Statistical Test: One-sample t-Test

Hypothesis Testing: H_0: No changes in number of women employed by local businesses; H_1: A change in number of women employed by local businesses

```
women_employees = ci_post_enc$workforce_change_female
hist(women_employees)
```




```
# One-sample t-Test
t.test(women_employees, mu = 0,alternative = "greater")
```

```
##
##  One Sample t-test
##
## data:  women_employees
## t = 7.6007, df = 369, p-value = 0.0000000000001229
## alternative hypothesis: true mean is greater than 0
## 95 percent confidence interval:
##  0.249729      Inf
## sample estimates:
## mean of x
## 0.3189189
```

Question 1-d: Are any female household members business owners? **Data set**: Household Paired Survey

Statistical Test: Paired Proportion McNemar's Chi-squared test

Hypothesis Testing: H_0: No changes in the proportion of women business owners; H_1: A change in the proportion of women business owners

```
business_women_matrix <- c(nrow(enc_paired %>%
                                  filter(business_owners_female_pre == 0 &
```

```
                                        business_owners_female_post == 0)),
                          nrow(enc_paired %>%
                                 filter(business_owners_female_pre == 0 &
                                        business_owners_female_post == 1)),
                          nrow(enc_paired %>%
                                 filter(business_owners_female_pre == 1 &
                                        business_owners_female_post == 0)),
                          nrow(enc_paired %>%
                                 filter(business_owners_female_pre == 1 &
                                        business_owners_female_post == 1))
)
Performance <-
    matrix(business_women_matrix,
           nrow = 2,
           dimnames = list("1st Survey" = c("No", "Yes"),
               "2nd Survey" = c("No", "Yes")))

Performance
```

```
##           2nd Survey
## 1st Survey  No Yes
##        No  573 226
##        Yes 144  67
```

```
# Approximation test
mcnemar.test(Performance)
```

```
##
##  McNemar's Chi-squared test with continuity correction
##
## data:  Performance
## McNemar's chi-squared = 17.732, df = 1, p-value = 0.00002543
```

```
## Exact test
mcnemar.exact(Performance)
```

```
##
##  Exact McNemar test (with central confidence intervals)
##
## data:  Performance
## b = 226, c = 144, p-value = 0.00002367
## alternative hypothesis: true odds ratio is not equal to 1
## 95 percent confidence interval:
##  1.267850 1.947937
## sample estimates:
## odds ratio
##   1.569444
```

2- PRODUCTIVITY

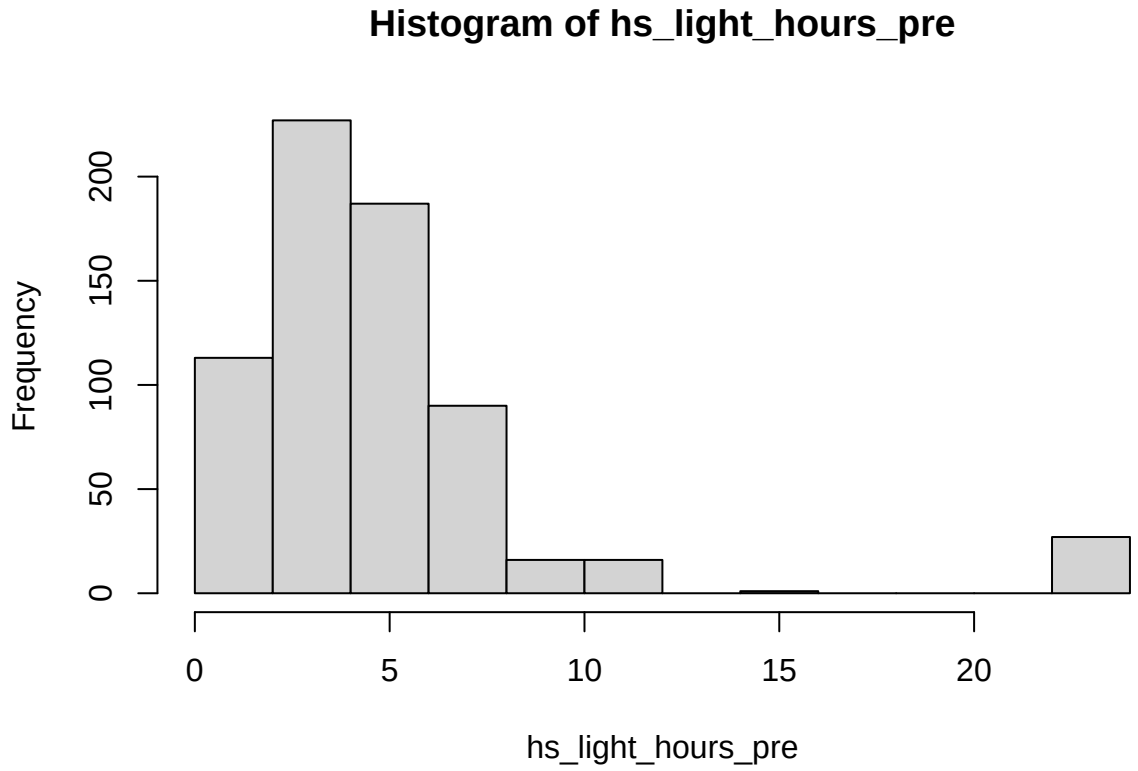
Question 2-a: How many hours of light per day do you currently have at home? Data set: Household Paired Survey

Statistical Test: Two-sample Paired t-Test

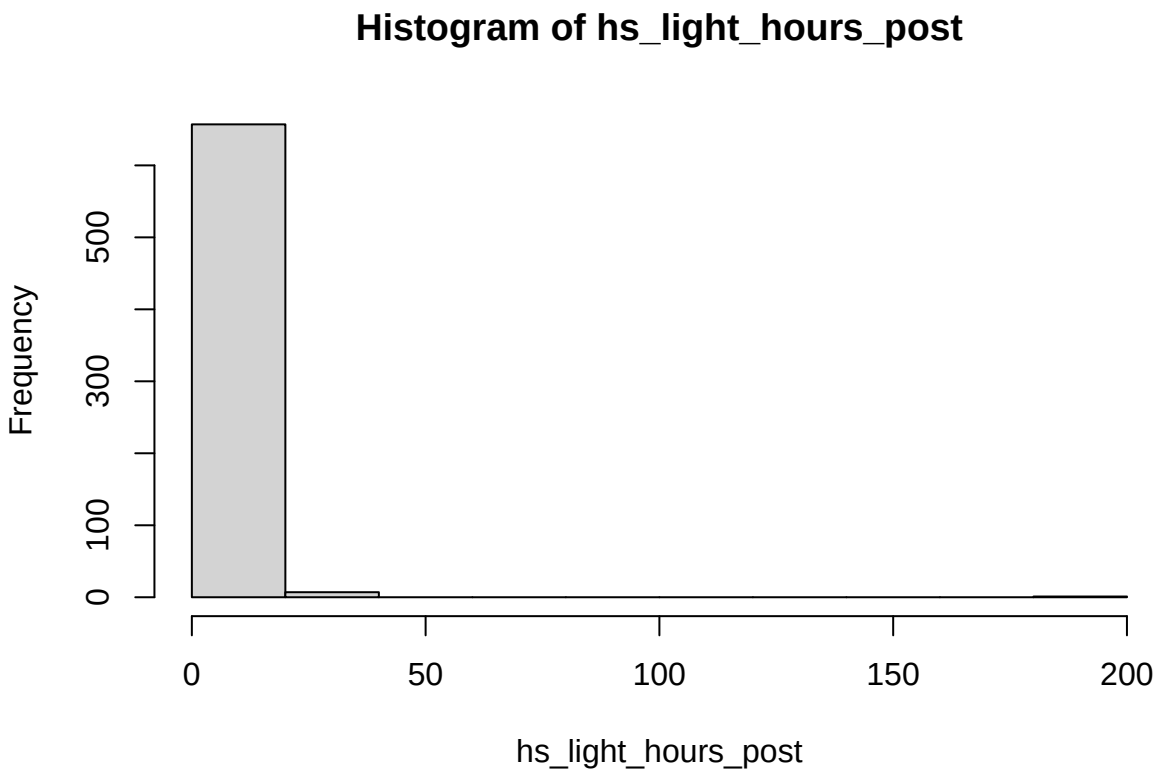
Hypothesis Testing: H_0: No change in the average light hours for household; H_1: A change in the average income for household

```
household <- enc_paired %>% filter(tariff_pre == "residential")
# commercial <- enc_paired %>% filter(tariff_pre == "commercial")

# For household
hs_light_hours_pre <- household$light_hours_current_pre
hist(hs_light_hours_pre)
```




```
hs_light_hours_post <- household$light_hours_current_post
hist(hs_light_hours_post)
```




```
t.test(hs_light_hours_pre, hs_light_hours_post,
       paired = TRUE, alternative = "two.sided")
```

```
##
##  Paired t-test
##
## data:  hs_light_hours_pre and hs_light_hours_post
## t = -5.1799, df = 588, p-value = 0.0000003054
## alternative hypothesis: true mean difference is not equal to 0
## 95 percent confidence interval:
##  -2.847295 -1.281737
## sample estimates:
## mean difference
##       -2.064516
```

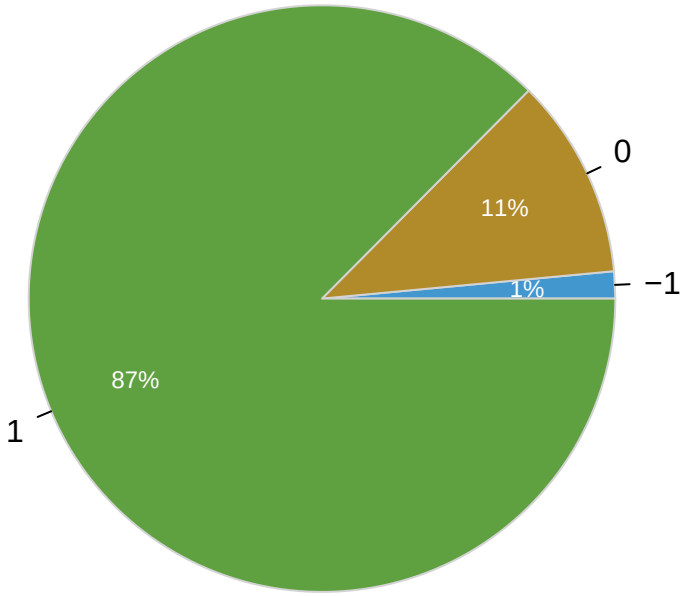
```
# For Commercial
# ci_light_hours_pre <- commercial$light_hours_current_pre
# ci_light_hours_post <- commercial$light_hours_current_post
# t.test(light_hours_pre, light_hours_post,
#        paired = TRUE, alternative = "two.sided")
```

Question 2-b: Since connection to Renewvia mini-grid, have your hours of operation changed at all? Data set: Commercial Post Survey

Statistical Test: Pearson's Chi-squared Goodness of fit test

Hypothesis Testing: H_0: No change in business operation hours; H_1: A change in business operation hours

```
# Breakdown of observed change in school enrollment of girls
PieChart(operations_hours_change, hole = 0, values = "%",
         data = ci_post_enc,
         main = "")
```




```
## >>> suggestions
## PieChart(operations_hours_change, hole=0)  # traditional pie chart
## PieChart(operations_hours_change, values="%")  # display %'s on the chart
## PieChart(operations_hours_change)  # bar chart
## Plot(operations_hours_change)  # bubble plot
## Plot(operations_hours_change, values="count")  # lollipop plot
##
## --- operations_hours_change ---
##
##                    -1      0      1     Total
## Frequencies:        7     52    411       470
## Proportions:    0.015  0.111  0.874     1.000
##
## Chi-squared test of null hypothesis of equal probabilities
##   Chisq = 625.791, df = 2, p-value = 0.000
```

```r
operation_hours <- factor(ci_post_enc$operations_hours_change)
operation_hours_table <-table(operation_hours)
chisq.test(operation_hours_table,p=rep(1/3,3))
```

```
##
##  Chi-squared test for given probabilities
##
## data:  operation_hours_table
## X-squared = 625.79, df = 2, p-value < 0.00000000000000022
```

Question 2-c: Since connection to Renewvia mini-grid, what appliances have you added to your business operations? Data set: Commercial Post Survey

Statistical Test: One sample t-Test

Hypothesis Testing: H_0: No change in number of appliances added to local business; H_1: A change in number of appliances added to local business

```r
# No of appliances
new_business_appliances <- ci_post_enc$appliances_addition_count

# One-sample t-Test
t.test(new_business_appliances, mu = 0,alternative = "greater")
```

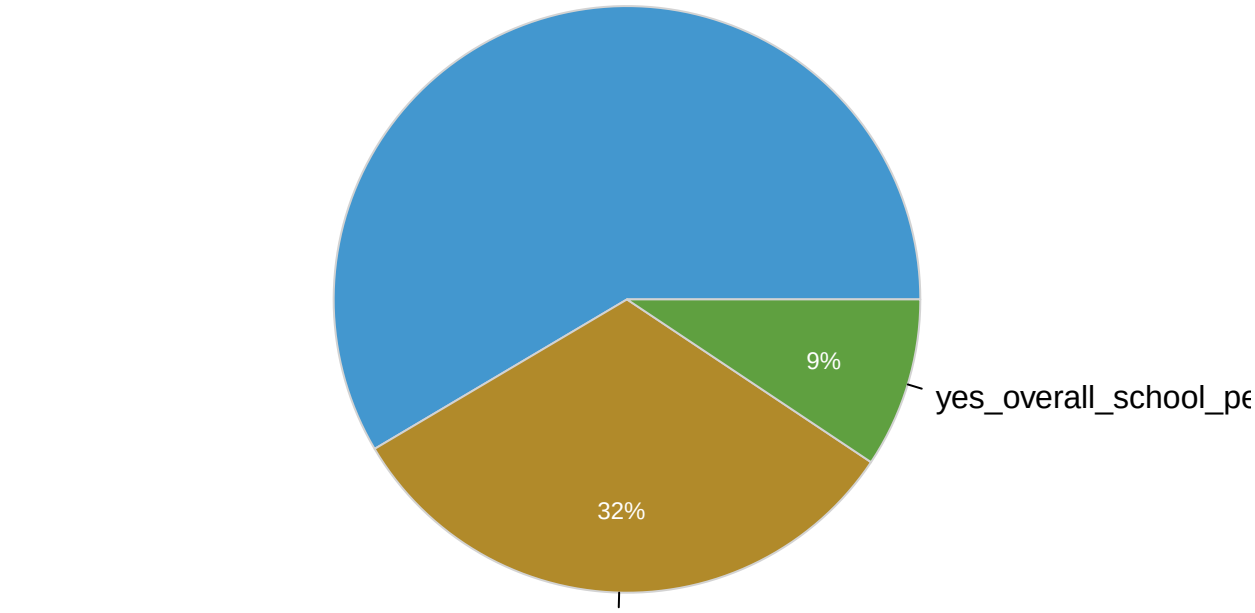
```
##
##  One Sample t-test
##
## data:  new_business_appliances
## t = 49.587, df = 469, p-value < 0.00000000000000022
## alternative hypothesis: true mean is greater than 0
## 95 percent confidence interval:
##  3.945217      Inf
## sample estimates:
## mean of x
##  4.080851
```

Question 2-c: Since connection to Renewvia mini-grid, what appliances have you added to your business operations? Data set: Commercial (Schools) Post Survey

Statistical Test: Pearon's Chi-Squared Test

Hypothesis Testing: H_0: No change in number of appliances added to local business; H_1: A change in number of appliances added to local business

```r
# Breakdown of observed change in school enrollment of girls
PieChart(school_performance, hole = 0, values = "%",
         data = ci_post_clean,
         main = "")
```




```
## >>> suggestions
## PieChart(school_performance, hole=0)  # traditional pie chart
## PieChart(school_performance, values="%")  # display %'s on the chart
## PieChart(school_performance)  # bar chart
## Plot(school_performance)  # bubble plot
## Plot(school_performance, values="count")  # lollipop plot
##
## --- school_performance ---
##
##                       n_thr_hs_nt_bn__ntc_  ys_vrll_schl_prfrm__     Total
## Frequencies:     275                   151                    44       470
## Proportions:   0.585                 0.321                 0.094     1.000
##
## Chi-squared test of null hypothesis of equal probabilities
##   Chisq = 170.609, df = 2, p-value = 0.000
```

```
school_performance <- factor(ci_post_enc$school_performance)
school_performance_table <-table(school_performance)
chisq.test(school_performance_table,p=rep(1/2,2))
```

```
##
##  Chi-squared test for given probabilities
##
## data:  school_performance_table
## X-squared = 58.713, df = 1, p-value = 0.00000000000001824
```

TO-DO

- Has school performance changed since connection to mini-grid?
- Have you seen a change in overall school performance?
- How many cell phones does your household have?
- How far must you travel to obtain your water supply?
- How long does the water collection process take on a daily basis?
- How long does the cooking fuel collection process take on a daily basis?
- Time spent traveling to charge phones
- How often do you need to charge your mobile phone?
- How far must you travel to charge your mobile phone?
- If you answered "yes, they have increased" above, by how many hours have they increased?

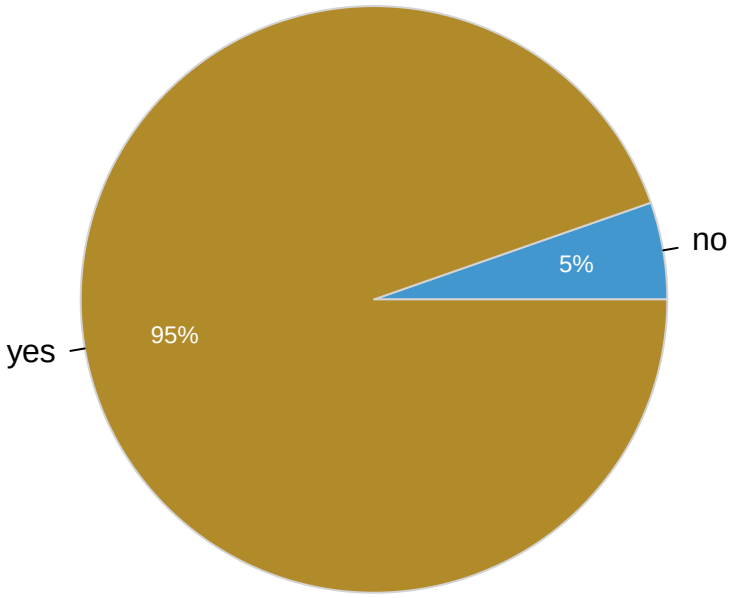
3- HEALTH

Question 3-a: Do you feel that you have better access to health services because of connection to mini-grid? Data set: Household Post Survey

Statistical Test: Pearson's Chi-Squared Test

Hypothesis Testing: H_0: No change in access to better healthcare; H_1: A change in access to better healthcare

```
# Breakdown of observed change
hs_post_clean_id <- hs_post_clean_id %>% filter(
                                                  better_access_health_minigrid != ""
                                                )
PieChart(better_access_health_minigrid, hole = 0, values = "%",
         data = hs_post_clean_id,
         main = "")
```




```
## >>> suggestions
## PieChart(better_access_health_minigrid, hole=0)  # traditional pie chart
## PieChart(better_access_health_minigrid, values="%")  # display %'s on the chart
## PieChart(better_access_health_minigrid)  # bar chart
## Plot(better_access_health_minigrid)  # bubble plot
## Plot(better_access_health_minigrid, values="count")  # lollipop plot
##
## --- better_access_health_minigrid ---
##
##                    no    yes      Total
## Frequencies:      115   2035       2150
## Proportions:    0.053  0.947      1.000
##
## Chi-squared test of null hypothesis of equal probabilities
##   Chisq = 1714.605, df = 1, p-value = 0.000
```

```
better_health <- factor(hs_post_clean_id$better_access_health_minigrid)
better_health_table <-table(better_health)
chisq.test(better_health_table,p=rep(1/2, 2))
```

```
##
##  Chi-squared test for given probabilities
##
## data:  better_health_table
## X-squared = 1714.6, df = 1, p-value < 0.00000000000000022
```

Question 3-B: Do you have a source for clean drinking water? **Data set**: Household Paired Survey

Statistical Test: McNemar's Chi-Squared Test

Hypothesis Testing: H_0: No change in access to better healthcare; H_1: A change in access to better healthcare

```
clean_water_matrix <- c(nrow(enc_paired %>%
                                 filter(clean_drinking_water_pre == 0 &
                                          clean_drinking_water_post == 0)),
                        nrow(enc_paired %>%
                                 filter(clean_drinking_water_pre == 0 &
                                          clean_drinking_water_post == 1)),
                        nrow(enc_paired %>%
                                 filter(clean_drinking_water_pre == 1 &
                                          clean_drinking_water_post == 0)),
                        nrow(enc_paired %>%
                                 filter(clean_drinking_water_pre == 1 &
                                          clean_drinking_water_post == 1))
)
Performance <-
    matrix(clean_water_matrix,
           nrow = 2,
           dimnames = list("1st Survey" = c("No", "Yes"),
               "2nd Survey" = c("No", "Yes")))

Performance
```

```
##           2nd Survey
## 1st Survey  No Yes
##        No  256  69
##        Yes 220 174
```

```
# Approximation test
mcnemar.test(Performance)
```

```
##
##  McNemar's Chi-squared test with continuity correction
##
## data:  Performance
## McNemar's chi-squared = 77.855, df = 1, p-value < 0.00000000000000022
```

```
## Exact test
mcnemar.exact(Performance)
```

```
##
##  Exact McNemar test (with central confidence intervals)
##
## data:  Performance
## b = 69, c = 220, p-value < 0.00000000000000022
## alternative hypothesis: true odds ratio is not equal to 1
## 95 percent confidence interval:
##  0.2357523 0.4127820
```

```
## sample estimates:
## odds ratio
##  0.3136364
```

- What hours is the Health Center / Clinic open?
- For clinics or health services only: Since connection to Renewvia minigrid, have your health service offerings changed in any of the following ways: (select all that apply).
- Does your Health Center / Clinic have access to refrigeration?
- Does your Health Center have access to minigrid electricity?
- What is the source for clean drinking water?
- How many kerosene lamps do you currently use in your household?
- Has the number of kerosene lamps in your household changed since connection to minigrid?
- Approximately how many hours a day do you use kerosene lamps?
- Do you feel that you have better access to health services because of connection to mini-grid?

4- SAFETY

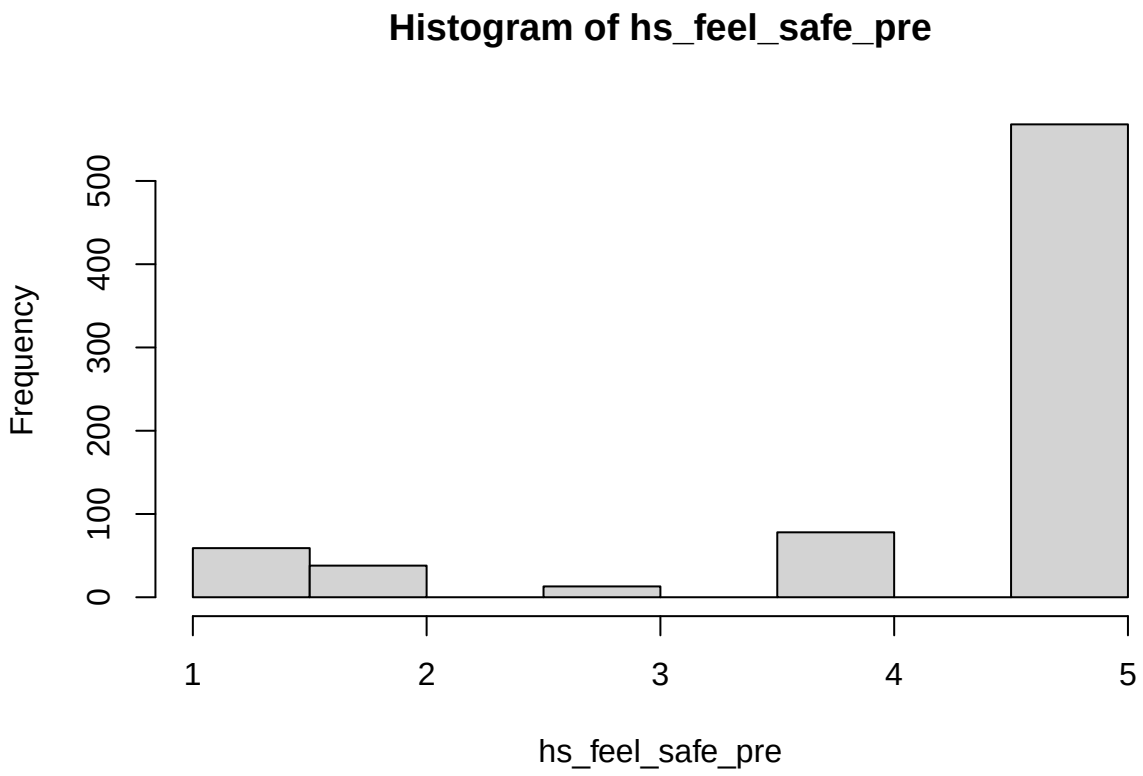
Question 4-a: How safe do you feel outside your home when it is dark? **Data set**: Household Paired Survey

Statistical Test: Paired t-Test

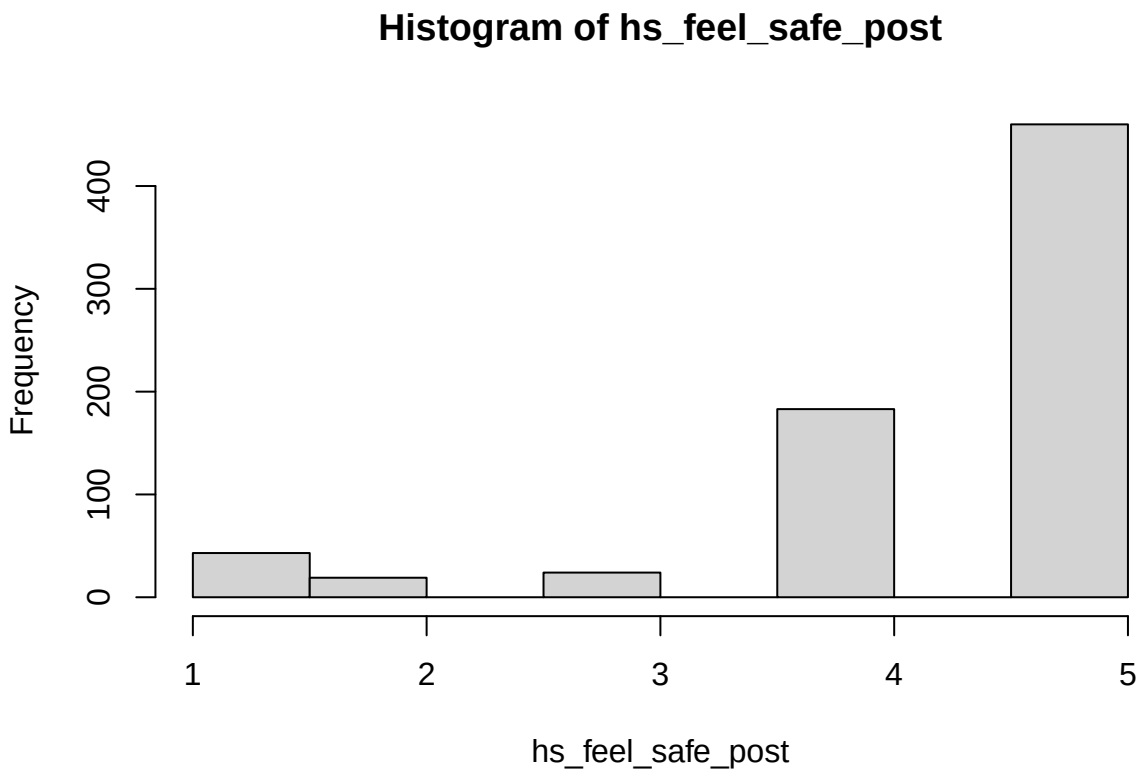
Hypothesis Testing: H_0: No change in feeling safe; H_1: A change in feeling safe

```
# household <- enc_paired %>% filter(tariff_pre == "residential")
# commercial <- enc_paired %>% filter(tariff_pre == "commercial")

# For household
hs_feel_safe_pre <- household$feel_safe_dark_pre
hist(hs_feel_safe_pre)
```




```r
hs_feel_safe_post <- household$feel_safe_dark_post
hist(hs_feel_safe_post)
```




```
t.test(hs_feel_safe_pre, hs_feel_safe_post,
       paired = TRUE, alternative = "two.sided")
```

```
##
##  Paired t-test
##
## data:  hs_feel_safe_pre and hs_feel_safe_post
## t = 0.08907, df = 724, p-value = 0.9291
## alternative hypothesis: true mean difference is not equal to 0
## 95 percent confidence interval:
##  -0.1160912  0.1271257
## sample estimates:
## mean difference
##     0.005517241
```

```
# For Commercial
# ci_light_hours_pre <- commercial$light_hours_current_pre
# ci_light_hours_post <- commercial$light_hours_current_post
# t.test(light_hours_pre, light_hours_post,
#        paired = TRUE, alternative = "two.sided")
```

TO-DO

- What makes you feel the most unsafe?

- Does your community currently have outdoor community lights?
- How safe would you feel outside your home at nighttime IF you had exterior lights?

5- ECONOMIC ACTIVITY

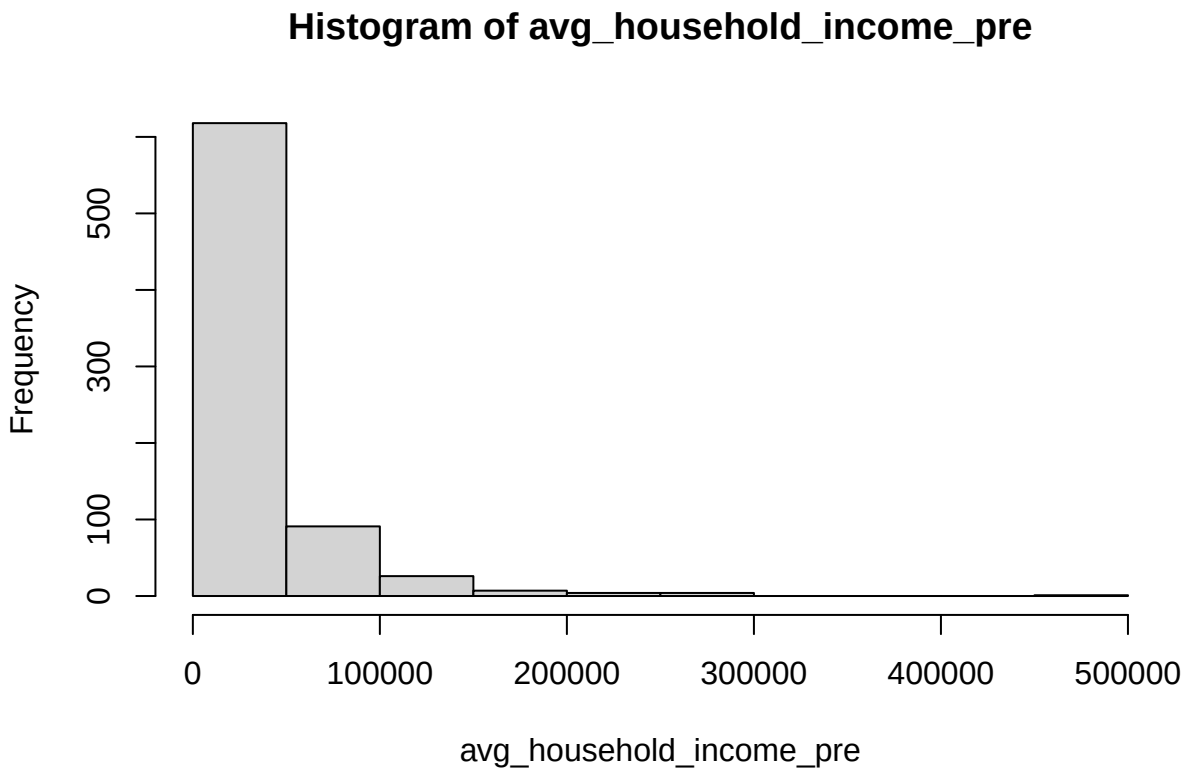
Question 5-a: What is your average monthly household income? **Data set**: Household Paired Survey

Statistical Test: Paired t-Test, setting mean to 3 as neutral

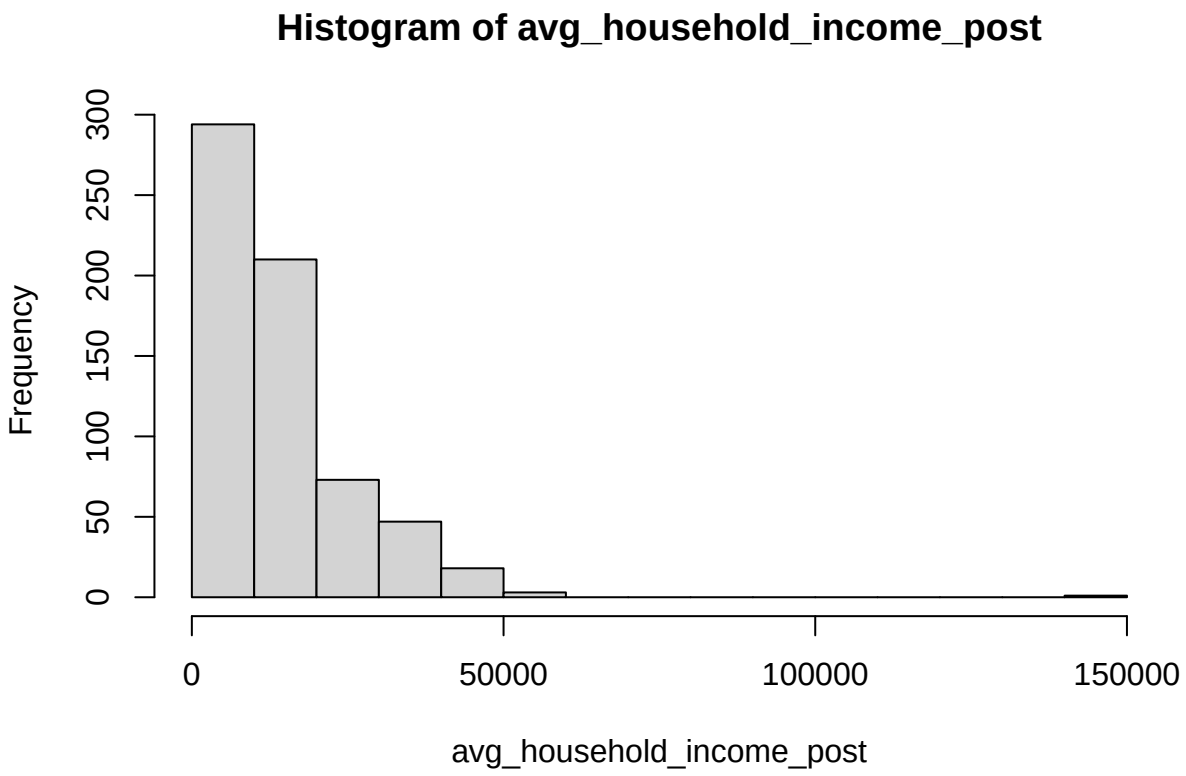
Hypothesis Testing: H_0: No change in average household income; H_1: A change in average household income

```
# household <- enc_paired %>% filter(tariff_pre == "residential")
# commercial <- enc_paired %>% filter(tariff_pre == "commercial")

# For household
avg_household_income_pre <- household$avg_household_income_pre
hist(avg_household_income_pre)
```




```
avg_household_income_post <- household$avg_household_income_post
hist(avg_household_income_post)
```




```
t.test(avg_household_income_pre, avg_household_income_post,
       paired = TRUE, alternative = "two.sided")
```

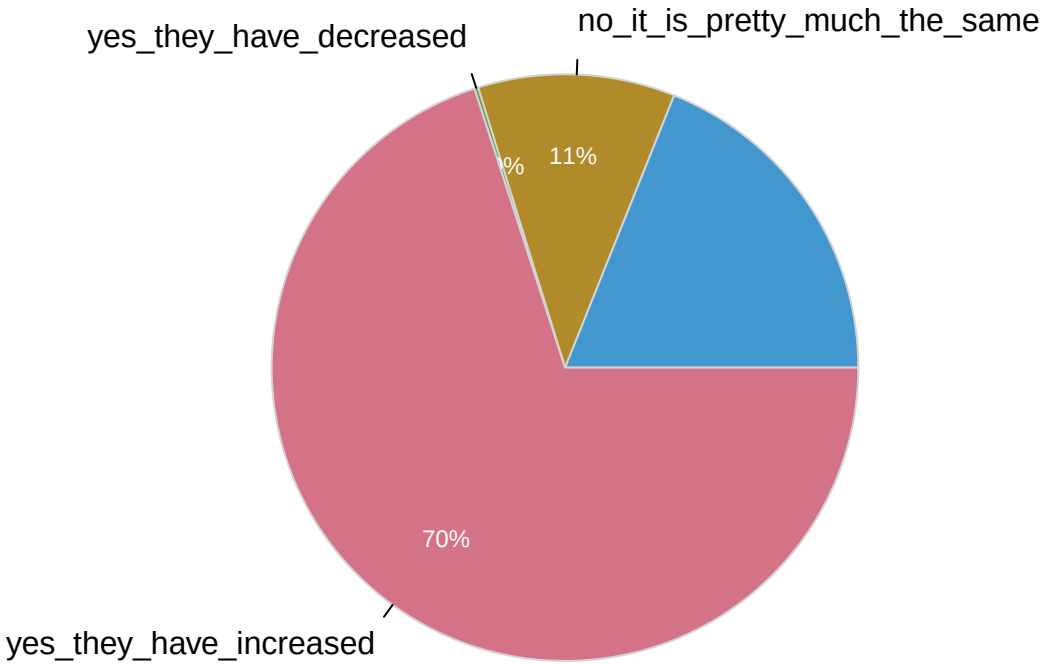
```
##
##  Paired t-test
##
## data:  avg_household_income_pre and avg_household_income_post
## t = 8.7187, df = 638, p-value < 0.00000000000000022
## alternative hypothesis: true mean difference is not equal to 0
## 95 percent confidence interval:
##  11971.98 18932.47
## sample estimates:
## mean difference
##        15452.23
```

Question 5-b: For businesses or shop owners only: Since connection to Renewvia mini-grid, have you seen any change in your weekly or monthly earnings? Data set: Commercial Post Survey

Statistical Test: Pearson's Chi-squared test

Hypothesis Testing: H_0: No change in earnings for business owners; H_1: A change in earnings for business owners

```r
# Breakdown of observed change
PieChart(earnings_change, hole = 0, values = "%",
         data = ci_post_clean,
         main = "")
```




```
## >>> suggestions
## PieChart(earnings_change, hole=0)  # traditional pie chart
## PieChart(earnings_change, values="%")  # display %'s on the chart
## PieChart(earnings_change)  # bar chart
## Plot(earnings_change)  # bubble plot
## Plot(earnings_change, values="count")  # lollipop plot
##
## --- earnings_change ---
##
##        earnings_chng Count   Prop
## ---------------------------------
##                           89   0.189
## n_t_s_prtty_mch_th_s      51   0.109
## yes_they_have_decrsd       1   0.002
## yes_they_have_incrsd     329   0.700
## ---------------------------------
##                  Total   470   1.000
##
## Chi-squared test of null hypothesis of equal probabilities
##   Chisq = 540.757, df = 3, p-value = 0.000
```

```r
earnings_change <- factor(ci_post_enc$earnings_change)
earnings_change_table <-table(earnings_change)
chisq.test(earnings_change_table,p=rep(1/3,3))
```

```
##
##  Chi-squared test for given probabilities
##
## data:  earnings_change_table
## X-squared = 491.78, df = 2, p-value < 0.00000000000000022
```

TO-DO

- Has your household income changed since your connection to minigrid power?
- What services are you able to offer/sell/provide due to connection to Renewvia minigrid that you weren't able to offer prior to connection?
- What is the Clinic / Hospital able to offer or provide due to connection to Renewvia minigrid that it wasn't able to offer before?
- What is your school able to offer / accomplish by having electricity that wasn't possible before?
- Are any household members business owners?
- Is this business recent (started after connection to mini-grid power?)
- If this business is new, do you consider this new business a result of having access to mini-grid power?
- Does this business use mini-grid power?
- Since connection to Renewvia minI-grid, have you had any change in number of workers/employees at your place of work?