Emily Chu Lab 5 Assignment

Code

This is an R Markdown (http://rmarkdown.rstudio.com) Notebook. When you execute code within the notebook, the results appear beneath the code.

Assignment 5:

Data frame for this assignment: country

1. Conduct the appropriate hypothesis test to determine if there is a statistically significant difference in female life expectancy 2004 based on HDI Category.

a. Use variables at the appropriate level of measurement.

HDI Category: Ordinal-level >> hdicat(x) Female life expectancy: Interval Ratio >>lifef(y)

b. Show R syntax.

aggregate(y~x, data, mean) anova.results <- aov(y~x, dataset)

c. Show results tables.

hdicat <chr>	lifef <dbl>
High Development	80.02909
Low Development	47.43750
Medium Development	68.09882

3 rows

Hide

```
summary(anova.hdicatLifef)
```

```
             Df Sum Sq Mean Sq F value Pr(>F)
hdicat        2  21489   10744   198.7 <2e-16 ***
Residuals   169   9140      54
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
21 observations deleted due to missingness
```

d. Report technical results.

For F=198.7 and pr(>F)=<2e-16 smaller than alpha 0.05, the result is statistically significant and we reject the null hypothesis.

e. In one paragraph, comment on the implications of the result of the test. Speculate. Is this relationship cause and effect; why or why not?

There is a statistically different difference in female life expectancy between low, medium and high countries. There are factors inclusive of hdicat, such as education, resources, access to care, economic opportunities that may create the effect of lower or higher female life expectancy between the different hdi categories. While the categorization itself may not be the direct cause, the factors leading to the country's categorizing impacts the female life expectancy.

2. If you wanted to conduct an ANOVA hypothesis test, but both of your variables were interval-ratio, which of the following tasks would you need to complete in order to be able to conduct the test? Select only one answer:

b. Recode the independent variable into a nominal or ordinal variable.