**Morgan Weyrens-Welch**

**Dr. Kristine West**

**Quantitative Program Evaluation**

**Problem Set 2**

**Friday, February 16th 2018**

> library(readxl)

> #AeonDataSheet1 <- read\_excel("Desktop/AeonData.xlsx")

> AeonDataSheet1 <- read\_excel("AeonData.xlsx")

Warning messages:

1: In read\_fun(path = path, sheet = sheet, limits = limits, shim = shim, :

Expecting numeric in K3179 / R3179C11: got 'N/A'

2: In read\_fun(path = path, sheet = sheet, limits = limits, shim = shim, :

Expecting numeric in K3271 / R3271C11: got 'N/A'

3: In read\_fun(path = path, sheet = sheet, limits = limits, shim = shim, :

Expecting numeric in K3303 / R3303C11: got 'N/A'

> View(AeonDataSheet1)

>

> positive<-as.numeric(AeonDataSheet1$'Exit Type'=="Positive")

>

> AeonDataSheet1<-cbind(AeonDataSheet1, positive)

> AeonDataSheet1noneutral<-subset(AeonDataSheet1,!(AeonDataSheet1$'Exit Type'=="Neutral"))

> t.test(AeonDataSheet1noneutral$'MI Income no subsidy'~AeonDataSheet1noneutral$`Exit Type`)

Welch Two Sample t-test

data: AeonDataSheet1noneutral$"MI Income no subsidy" by AeonDataSheet1noneutral$`Exit Type`

t = -9.3442, df = 1705.4, p-value < 2.2e-16

alternative hypothesis: true difference in means is not equal to 0

95 percent confidence interval:

-7845.036 -5123.024

sample estimates:

mean in group Negative mean in group Positive

18264.87 24748.90

>

> #without neutrals

> mean(AeonDataSheet1noneutral$'MI Income no subsidy',na.rm=TRUE)

[1] 22961.82

>

> # total mean with neutrals

> mean(AeonDataSheet1$'MI Income no subsidy',na.rm=TRUE)

[1] 20920.48

>

> #Correlation 2a

> library("car", lib.loc="/Library/Frameworks/R.framework/Versions/3.4/Resources/library")

> scatterplot(AeonDataSheet1$'Late Rent Count'~AeonDataSheet1$'MI Income no subsidy', data=AeonDataSheet1, xlab="MI Income no subsidy", ylab="Late Rent Count", main="MI Income and Late Rent Count")

>

> #Regression without dummies

> lm(AeonDataSheet1$`Late Rent Count`~AeonDataSheet1$'MI Income no subsidy'+ AeonDataSheet1$'First Subsidy Amt')

Call:

lm(formula = AeonDataSheet1$`Late Rent Count` ~ AeonDataSheet1$"MI Income no subsidy" +

AeonDataSheet1$"First Subsidy Amt")

Coefficients:

(Intercept) AeonDataSheet1$"MI Income no subsidy" AeonDataSheet1$"First Subsidy Amt"

6.019e+00 -3.459e-05 -2.851e-03

> model2<- lm(AeonDataSheet1$`Late Rent Count`~AeonDataSheet1$'MI Income no subsidy'+ AeonDataSheet1$'First Subsidy Amt')

> ##dummies

> socialsecurity<-as.numeric(AeonDataSheet1$'Primary Income Source'=="Social Security")

> fedwage<-as.numeric(AeonDataSheet1$'Primary Income Source'=="Federal Wage")

> genassistance<-as.numeric(AeonDataSheet1$'Primary Income Source'=="General Assistance")

> othernonwage<-as.numeric(AeonDataSheet1$'Primary Income Source'=="Other Nonwage Source")

> nonfederalwage<-as.numeric(AeonDataSheet1$'Primary Income Source'=="Nonfederal Wage")

> ssi<-as.numeric(AeonDataSheet1$'Primary Income Source'=="SSI")

> unemployment<-as.numeric(AeonDataSheet1$'Primary Income Source'=="Unemployment")

> tanif<-as.numeric(AeonDataSheet1$'Primary Income Source'=="TANIF")

> business<-as.numeric(AeonDataSheet1$'Primary Income Source'=="Business")

> childsupport<-as.numeric(AeonDataSheet1$'Primary Income Source'=="Child Support")

> pensions<-as.numeric(AeonDataSheet1$'Primary Income Source'=="Pensions")

> militarypay<-as.numeric(AeonDataSheet1$'Primary Income Source'=="Military Pay")

>

> AeonDataSheet1<-cbind(AeonDataSheet1,business, childsupport, fedwage, genassistance, militarypay, nonfederalwage, othernonwage, pensions, positive, socialsecurity, ssi, tanif, unemployment)

> lm(AeonDataSheet1$`Late Rent Count`~AeonDataSheet1$'MI Income no subsidy'+ AeonDataSheet1$'First Subsidy Amt'

+ +AeonDataSheet1$socialsecurity+AeonDataSheet1$fedwage+AeonDataSheet1$genassistance+AeonDataSheet1$othernonwage

+ +AeonDataSheet1$nonfederalwage+AeonDataSheet1$ssi+AeonDataSheet1$unemployment+AeonDataSheet1$tanif+AeonDataSheet1$business

+ +AeonDataSheet1$childsupport+AeonDataSheet1$pensions+AeonDataSheet1$militarypay)

Call:

lm(formula = AeonDataSheet1$`Late Rent Count` ~ AeonDataSheet1$"MI Income no subsidy" +

AeonDataSheet1$"First Subsidy Amt" + AeonDataSheet1$socialsecurity +

AeonDataSheet1$fedwage + AeonDataSheet1$genassistance + AeonDataSheet1$othernonwage +

AeonDataSheet1$nonfederalwage + AeonDataSheet1$ssi + AeonDataSheet1$unemployment +

AeonDataSheet1$tanif + AeonDataSheet1$business + AeonDataSheet1$childsupport +

AeonDataSheet1$pensions + AeonDataSheet1$militarypay)

Coefficients:

(Intercept) AeonDataSheet1$"MI Income no subsidy" AeonDataSheet1$"First Subsidy Amt"

2.693e+01 -1.033e-04 -3.101e-03

AeonDataSheet1$socialsecurity AeonDataSheet1$fedwage AeonDataSheet1$genassistance

-2.109e+01 -1.145e+01 -2.218e+01

AeonDataSheet1$othernonwage AeonDataSheet1$nonfederalwage AeonDataSheet1$ssi

-2.135e+01 -1.744e+01 -2.034e+01

AeonDataSheet1$unemployment AeonDataSheet1$tanif AeonDataSheet1$business

-1.853e+01 -1.862e+01 -1.922e+01

AeonDataSheet1$childsupport AeonDataSheet1$pensions AeonDataSheet1$militarypay

-1.713e+01 -2.448e+01 NA

>

>

> model1<-lm(AeonDataSheet1$`Late Rent Count`~AeonDataSheet1$'MI Income no subsidy'+ AeonDataSheet1$'First Subsidy Amt'

+ +AeonDataSheet1$socialsecurity+AeonDataSheet1$fedwage+AeonDataSheet1$genassistance+AeonDataSheet1$othernonwage

+ +AeonDataSheet1$nonfederalwage+AeonDataSheet1$ssi+AeonDataSheet1$unemployment+AeonDataSheet1$tanif+AeonDataSheet1$business

+ +AeonDataSheet1$childsupport+AeonDataSheet1$pensions+AeonDataSheet1$militarypay)

>

> summary(model2)

Call:

lm(formula = AeonDataSheet1$`Late Rent Count` ~ AeonDataSheet1$"MI Income no subsidy" +

AeonDataSheet1$"First Subsidy Amt")

Residuals:

Min 1Q Median 3Q Max

-6.019 -4.371 -2.774 1.464 69.318

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) 6.019e+00 2.653e-01 22.690 < 2e-16 \*\*\*

AeonDataSheet1$"MI Income no subsidy" -3.459e-05 8.755e-06 -3.951 7.93e-05 \*\*\*

AeonDataSheet1$"First Subsidy Amt" -2.851e-03 4.558e-04 -6.256 4.43e-10 \*\*\*

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 7.659 on 3545 degrees of freedom

(224 observations deleted due to missingness)

Multiple R-squared: 0.01159, Adjusted R-squared: 0.01103

F-statistic: 20.79 on 2 and 3545 DF, p-value: 1.057e-09

> summary(model1)

Call:

lm(formula = AeonDataSheet1$`Late Rent Count` ~ AeonDataSheet1$"MI Income no subsidy" +

AeonDataSheet1$"First Subsidy Amt" + AeonDataSheet1$socialsecurity +

AeonDataSheet1$fedwage + AeonDataSheet1$genassistance + AeonDataSheet1$othernonwage +

AeonDataSheet1$nonfederalwage + AeonDataSheet1$ssi + AeonDataSheet1$unemployment +

AeonDataSheet1$tanif + AeonDataSheet1$business + AeonDataSheet1$childsupport +

AeonDataSheet1$pensions + AeonDataSheet1$militarypay)

Residuals:

Min 1Q Median 3Q Max

-13.691 -4.427 -2.306 1.627 53.793

Coefficients: (1 not defined because of singularities)

Estimate Std. Error t value Pr(>|t|)

(Intercept) 2.693e+01 7.896e+00 3.410 0.00066 \*\*\*

AeonDataSheet1$"MI Income no subsidy" -1.033e-04 2.222e-05 -4.649 3.51e-06 \*\*\*

AeonDataSheet1$"First Subsidy Amt" -3.101e-03 5.826e-04 -5.323 1.11e-07 \*\*\*

AeonDataSheet1$socialsecurity -2.109e+01 7.897e+00 -2.671 0.00760 \*\*

AeonDataSheet1$fedwage -1.145e+01 7.991e+00 -1.433 0.15197

AeonDataSheet1$genassistance -2.218e+01 7.913e+00 -2.804 0.00509 \*\*

AeonDataSheet1$othernonwage -2.135e+01 7.918e+00 -2.696 0.00706 \*\*

AeonDataSheet1$nonfederalwage -1.744e+01 7.889e+00 -2.210 0.02718 \*

AeonDataSheet1$ssi -2.034e+01 7.903e+00 -2.573 0.01013 \*

AeonDataSheet1$unemployment -1.853e+01 8.031e+00 -2.307 0.02114 \*

AeonDataSheet1$tanif -1.862e+01 8.071e+00 -2.306 0.02116 \*

AeonDataSheet1$business -1.922e+01 7.954e+00 -2.416 0.01574 \*

AeonDataSheet1$childsupport -1.713e+01 8.033e+00 -2.133 0.03306 \*

AeonDataSheet1$pensions -2.448e+01 8.430e+00 -2.904 0.00371 \*\*

AeonDataSheet1$militarypay NA NA NA NA

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 7.885 on 2619 degrees of freedom

(1139 observations deleted due to missingness)

Multiple R-squared: 0.07068, Adjusted R-squared: 0.06607

F-statistic: 15.32 on 13 and 2619 DF, p-value: < 2.2e-16