# GV300 - Quantitative Political Analysis

University of Essex - Department of Government

Lorenzo Crippa

Week 10 – 2 December, 2019

# Regression Analysis

Today we'll work on regression analysis and apply what you have learnt in class. We'll use the dataset on sales of Monet paintings that we have started to explore in week 8.

# Explaining prices of paintings – Part 1

Load your data of Monet paintings and perform the following steps:

1. Explore how variables relate to each other (bivariate *descriptive* analysis)
   - Use appropriate statistics to talk about these relations
   - Use appropriate plots to show the relations
2. Propose a theory about what explains the price of a painting
   - What is your theory?
   - What is the underlining population of your theory?
   - What is the sample you are using?
3. Build an appropriate linear model to test your theory
   - What variables do you include?
   - Think and justify whether you meet Gauss-Markov assumptions
   - What could violate them?
   - What could improve your model?

# Explaining prices of paintings – Part 2

4. Run your model using your preferred statistical software
    - Interpret the coefficients
    - What test is performed on each variable?
    - What variables are significant? Is your theory rejected?
    - Do you see any problem in your model?
5. How good is your model, jointly considered?
    - What test is performed on the overall model?
    - Interpret the F-test
    - Interpret the $R^2$
    - Which one of the two would you use to evaluate your model?
6. Based on your model, predict what would be the price of a painting with arbitrary values for the dependent variables

## Exercise 1

1. Explore how variables relate to each other. Appropriate statistics

# Exercise 1

1. Explore how variables relate to each other. Appropriate statistics

In R:

```r
cor(x = matrix(c(Greene$PRICE, Greene$HEIGHT,
         Greene$WIDTH), ncol = 3, byrow = FALSE),
    use = "pairwise.complete.obs")
```

## Exercise 1

1. Explore how variables relate to each other. Appropriate statistics

In R:

```r
cor(x = matrix(c(Greene$PRICE, Greene$HEIGHT,
         Greene$WIDTH), ncol = 3, byrow = FALSE),
    use = "pairwise.complete.obs")
```

Outputs:

```
          [,1]      [,2]      [,3]
[1,] 1.0000000 0.3145808 0.3468806
[2,] 0.3145808 1.0000000 0.5032801
[3,] 0.3468806 0.5032801 1.0000000
```

# Exercise 1

1. Explore how variables relate to each other. Appropriate statistics

# Exercise 1

1. Explore how variables relate to each other. Appropriate statistics

In Stata:

```
pwcorr price height width
```

## Exercise 1

1. Explore how variables relate to each other. Appropriate statistics

In Stata:

```
pwcorr price height width
```

Outputs:

```
             |    price   height    width
-------------+---------------------------
       price |   1.0000
      height |   0.3146   1.0000
       width |   0.3469   0.5033   1.0000
```

# Exercise 1

1. Explore how variables relate to each other. Appropriate plots

# Exercise 1

1. Explore how variables relate to each other. Appropriate plots

In R:

```
# scatterplots (for continuous variables)
ggplot(Greene, aes(x = HEIGHT, y = PRICE)) + geom_
    point() + xlab("height") + ylab("price")

# boxplots (for ordinal variables)
ggplot(Greene, aes(x = SIGNED, y = PRICE)) + geom_
    boxplot() + xlab("signed") + ylab("price")

# multivariate relations
ggplot(Greene, aes(x = WIDTH, y = PRICE, col = SIGNED)
    ) + geom_point() +
  xlab("width") + ylab("price") +
  scale_color_discrete("signed", breaks = c(0,1),
                       labels = c("no", "yes"))
```

# Exercise 1

1. Explore how variables relate to each other. Appropriate plots

1. Explore how variables relate to each other. Appropriate plots

In Stata:

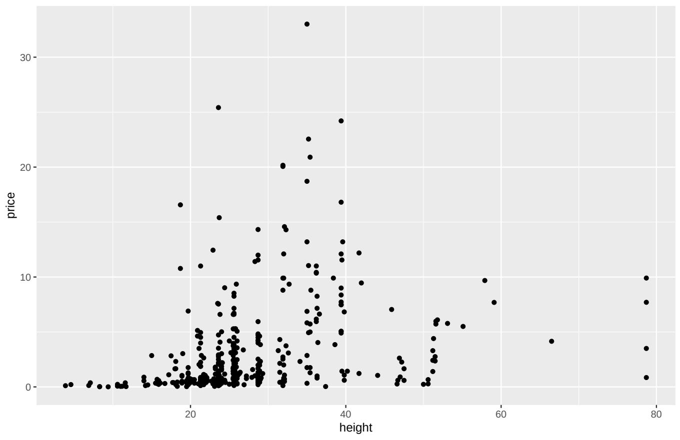
```
* scatterplots
twoway scatter price height

* boxplots
graph box price, over(signed)

* multivariate
twoway (scatter price height if signed == 0) ///
  (scatter price height if signed == 1), ///
  legend(label(1 not signed) label(2 signed))
```

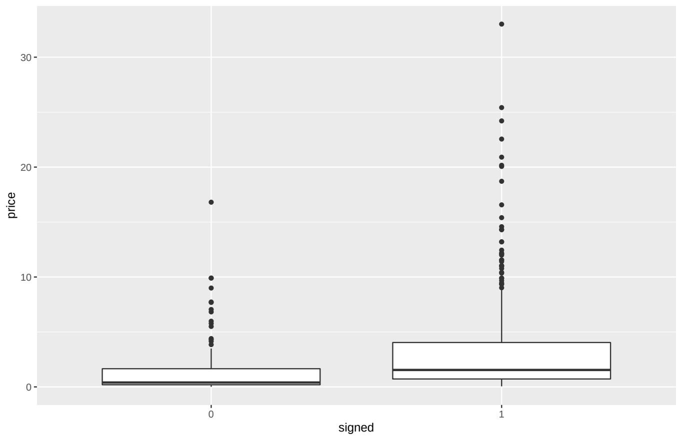
# Exercise 1

Scatterplots:



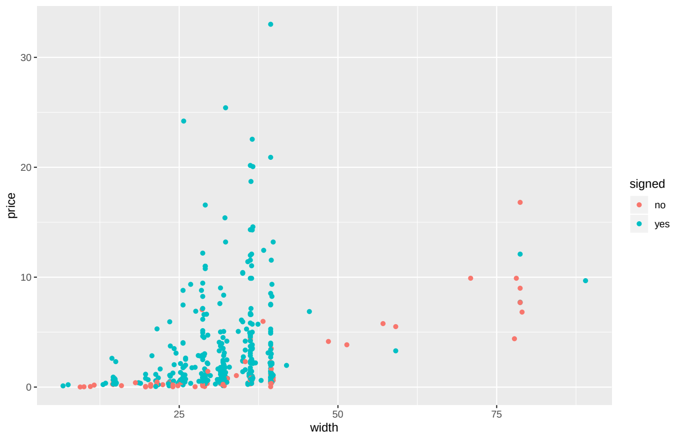
# Exercise 1

Boxplots:



# Exercise 1

Multivariate:



## Exercise 2

2. Propose a theory about what explains the price of a painting

# Exercise 2

2. Propose a theory about what explains the price of a painting

**Theory**

The price of a painting from Monet is explained (determined) by its dimensions: the larger the dimension, the higher the price.

# Exercise 2

2. Propose a theory about what explains the price of a painting

**Theory**

The price of a painting from Monet is explained (determined) by its dimensions: the larger the dimension, the higher the price.

**Population**

The population is represented by the sales of all paintings from Monet.

# Exercise 2

2. Propose a theory about what explains the price of a painting

**Theory**

The price of a painting from Monet is explained (determined) by its dimensions: the larger the dimension, the higher the price.

**Population**

The population is represented by the sales of all paintings from Monet.

**Sample**

The sample observations are sales of paintings from Monet from houses number 1, 2, 3.

# Exercise 3

3. Build an appropriate linear model to test your theory

## Exercise 3

3. Build an appropriate linear model to test your theory

$$Price_i = \beta_0 + \beta_1 Height_i + \beta_2 Width_i + \beta_3 Signed_i + \beta_4 House_i + u_i$$

# Exercise 3

3. Build an appropriate linear model to test your theory

$$Price_i = \beta_0 + \beta_1 Height_i + \beta_2 Width_i + \beta_3 Signed_i + \beta_4 House_i + u_i$$

Gauss-Markov assumptions make OLS BLUE (Greene, 2003):

## Exercise 3

3. Build an appropriate linear model to test your theory

$$Price_i = \beta_0 + \beta_1 Height_i + \beta_2 Width_i + \beta_3 Signed_i + \beta_4 House_i + u_i$$

Gauss-Markov assumptions make OLS BLUE (Greene, 2003):

1. Linearity

## Exercise 3

3. Build an appropriate linear model to test your theory

$$Price_i = \beta_0 + \beta_1 Height_i + \beta_2 Width_i + \beta_3 Signed_i + \beta_4 House_i + u_i$$

Gauss-Markov assumptions make OLS BLUE (Greene, 2003):

1. Linearity
2. No *perfect* multicollinearity

## Exercise 3

3. Build an appropriate linear model to test your theory

$$Price_i = \beta_0 + \beta_1 Height_i + \beta_2 Width_i + \beta_3 Signed_i + \beta_4 House_i + u_i$$

Gauss-Markov assumptions make OLS BLUE (Greene, 2003):

1. Linearity
2. No *perfect* multicollinearity
3. Exogeneity of independent variables, zero-conditional mean (no confounder, no OVB)

# Exercise 3

3. Build an appropriate linear model to test your theory

$$Price_i = \beta_0 + \beta_1 Height_i + \beta_2 Width_i + \beta_3 Signed_i + \beta_4 House_i + u_i$$

Gauss-Markov assumptions make OLS BLUE (Greene, 2003):

1. Linearity
2. No *perfect* multicollinearity
3. Exogeneity of independent variables, zero-conditional mean (no confounder, no OVB) $E(u_i|X_1, \dots, X_k) = 0$

# Exercise 3

3. Build an appropriate linear model to test your theory

$$Price_i = \beta_0 + \beta_1 Height_i + \beta_2 Width_i + \beta_3 Signed_i + \beta_4 House_i + u_i$$

Gauss-Markov assumptions make OLS BLUE (Greene, 2003):

1. Linearity
2. No *perfect* multicollinearity
3. Exogeneity of independent variables, zero-conditional mean (no confounder, no OVB) $E(u_i|X_1, \ldots, X_k) = 0$
4. Observations are i.i.d.

## Exercise 3

3. Build an appropriate linear model to test your theory

$$Price_i = \beta_0 + \beta_1 Height_i + \beta_2 Width_i + \beta_3 Signed_i + \beta_4 House_i + u_i$$

Gauss-Markov assumptions make OLS BLUE (Greene, 2003):

1. Linearity
2. No *perfect* multicollinearity
3. Exogeneity of independent variables, zero-conditional mean (no confounder, no OVB) $E(u_i|X_1, \dots, X_k) = 0$
4. Observations are i.i.d.
5. Constant variance of the *error* term* (homoskedasticity)

## Exercise 3

3. Build an appropriate linear model to test your theory

$$Price_i = \beta_0 + \beta_1 Height_i + \beta_2 Width_i + \beta_3 Signed_i + \beta_4 House_i + u_i$$

Gauss-Markov assumptions make OLS BLUE (Greene, 2003):

1. Linearity
2. No *perfect* multicollinearity
3. Exogeneity of independent variables, zero-conditional mean (no confounder, no OVB) $E(u_i|X_1, \ldots, X_k) = 0$
4. Observations are i.i.d.
5. Constant variance of the *error* term* (homoskedasticity)
6. Normal distribution of the *error* term*

# Exercise 3

3. Build an appropriate linear model to test your theory

$$Price_i = \beta_0 + \beta_1 Height_i + \beta_2 Width_i + \beta_3 Signed_i + \beta_4 House_i + u_i$$

Gauss-Markov assumptions make OLS BLUE (Greene, 2003):

1. Linearity
2. No *perfect* multicollinearity
3. Exogeneity of independent variables, zero-conditional mean (no confounder, no OVB) $E(u_i|X_1, \ldots, X_k) = 0$
4. Observations are i.i.d.
5. Constant variance of the *error* term* (homoskedasticity)
6. Normal distribution of the *error* term*

I could improve my model by including $Width \times Height$

## Exercise 3

3. Build an appropriate linear model to test your theory

$$Price_i = \beta_0 + \beta_1 Height_i + \beta_2 Width_i + \beta_3 Signed_i + \beta_4 House_i + u_i$$

Gauss-Markov assumptions make OLS BLUE (Greene, 2003):

1. Linearity
2. No *perfect* multicollinearity
3. Exogeneity of independent variables, zero-conditional mean (no confounder, no OVB) $E(u_i|X_1, \ldots, X_k) = 0$
4. Observations are i.i.d.
5. Constant variance of the *error* term* (homoskedasticity)
6. Normal distribution of the *error* term*

I could improve my model by including $Width \times Height$

**not* necessary

# Exercise 4

4. Run your model using your preferred statistical software

# Exercise 4

4. Run your model using your preferred statistical software

In R:

```
model <- lm(data = Greene, PRICE ~ HEIGHT + WIDTH +
    SIGNED + HOUSE)
summary(model)
```

# Exercise 4

4. Run your model using your preferred statistical software

In R:

```
model <- lm(data = Greene, PRICE ~ HEIGHT + WIDTH +
    SIGNED + HOUSE)
summary(model)
```

In Stata:

```
reg price height width signed house
```

# Exercise 4

OLS can also be computed using matrix algebra:

# Exercise 4

OLS can also be computed using matrix algebra:

- $\mathbf{X}$ is the $n \times k$ matrix of $k - 1$ independent variables (the intercept is the $+1$ variable)

## Exercise 4

OLS can also be computed using matrix algebra:

- $\mathbf{X}$ is the $n \times k$ matrix of $k - 1$ independent variables (the intercept is the $+1$ variable)
- $\mathbf{y}$ is the $n \times 1$ vector with observations of the dependent variable

## Exercise 4

OLS can also be computed using matrix algebra:

- $\mathbf{X}$ is the $n \times k$ matrix of $k-1$ independent variables (the intercept is the $+1$ variable)
- $\mathbf{y}$ is the $n \times 1$ vector with observations of the dependent variable
- $\mathbf{A}'$ is the transpose matrix of $\mathbf{A}$ (see math refresher)

## Exercise 4

OLS can also be computed using matrix algebra:

- $\mathbf{X}$ is the $n \times k$ matrix of $k-1$ independent variables (the intercept is the $+1$ variable)
- $\mathbf{y}$ is the $n \times 1$ vector with observations of the dependent variable
- $\mathbf{A}'$ is the transpose matrix of $\mathbf{A}$ (see math refresher)
- $\mathbf{A}^{-1}$ is the inverse of $\mathbf{A}$: $\mathbf{A} \times \mathbf{A}^{-1} = \mathbf{I}$ (see math refresher)

# Exercise 4

OLS can also be computed using matrix algebra:

- $\mathbf{X}$ is the $n \times k$ matrix of $k-1$ independent variables (the intercept is the $+1$ variable)
- $\mathbf{y}$ is the $n \times 1$ vector with observations of the dependent variable
- $\mathbf{A}'$ is the transpose matrix of $\mathbf{A}$ (see math refresher)
- $\mathbf{A}^{-1}$ is the inverse of $\mathbf{A}$: $\mathbf{A} \times \mathbf{A}^{-1} = \mathbf{I}$ (see math refresher)

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y} \tag{1}$$

# Exercise 4

Computing OLS using matrix algebra (see R script):

## Exercise 4

Computing OLS using matrix algebra (see R script):

```
Y <- Greene$PRICE

X <- matrix(c(rep(1, length(Greene$PRICE)),
              Greene$HEIGHT, Greene$WIDTH,
              Greene$SIGNED, Greene$HOUSE),
            ncol = 5, byrow = F)

OLS <- solve(t(X)%*%X) %*% (t(X) %*% Y)
```

# Exercise 4

Computing OLS using matrix algebra (see R script):

```r
Y <- Greene$PRICE

X <- matrix(c(rep(1, length(Greene$PRICE)),
              Greene$HEIGHT, Greene$WIDTH,
              Greene$SIGNED, Greene$HOUSE),
            ncol = 5, byrow = F)

OLS <- solve(t(X)%*%X) %*% (t(X) %*% Y)
```

Outputs:

# Exercise 4

Computing OLS using matrix algebra (see R script):

```r
Y <- Greene$PRICE

X <- matrix(c(rep(1, length(Greene$PRICE)),
              Greene$HEIGHT, Greene$WIDTH,
              Greene$SIGNED, Greene$HOUSE),
            ncol = 5, byrow = F)

OLS <- solve(t(X)%*%X) %*% (t(X) %*% Y)
```

Outputs:

```
             [,1]
[1,] -5.52043300
[2,]  0.09096657
[3,]  0.11169822
[4,]  2.29231727
[5,]  0.38896084
```

# Exercise 4

Results from the `stargazer` package in R (Hlavac, 2018):

| | *Dependent variable:* |
|---|---|
| | PRICE |
| HEIGHT | 0.091*** |
| | (0.022) |
| WIDTH | 0.112*** |
| | (0.021) |
| SIGNED | 2.292*** |
| | (0.503) |
| HOUSE | 0.389 |
| | (0.328) |
| Constant | −5.520*** |
| | (1.092) |
| Observations | 430 |
| Adjusted $R^2$ | 0.179 |
| F Statistic | 24.395*** (df = 4; 425) |
| *Note:* | *p<0.1; **p<0.05; ***p<0.01 |

## Exercise 4

Interpretation of the model:

# Exercise 4

Interpretation of the model:

$$\begin{aligned} Price = &-5.520 + 0.091 Height + 0.112 Width \\ &+ 2.292 Signed + 0.389 House + u \end{aligned} \tag{2}$$

## Exercise 4

Interpretation of the model:

$$\begin{aligned} \mathit{Price} = & -5.520 + 0.091\mathit{Height} + 0.112\mathit{Width} \\ & +2.292\mathit{Signed} + 0.389\mathit{House} + u \end{aligned} \tag{2}$$

- *Ceteris paribus*, 1 inch increase in *Height* (*Width*) raises the price by 0.091 (0.112) units. Similar for the other variables.

## Exercise 4

Interpretation of the model:

$$
\begin{aligned}
Price = &-5.520 + 0.091 Height + 0.112 Width \\
&+ 2.292 Signed + 0.389 House + u
\end{aligned}
\tag{2}
$$

- *Ceteris paribus*, 1 inch increase in *Height* (*Width*) raises the price by 0.091 (0.112) units. Similar for the other variables.
- For each variable the t-test is relative to the null-hypothesis that the true parameter (population parameter) equals 0.

# Exercise 4

Interpretation of the model:

$$\begin{aligned} Price = -5.520 + 0.091 Height + 0.112 Width \\ +2.292 Signed + 0.389 House + u \end{aligned} \tag{2}$$

- *Ceteris paribus*, 1 inch increase in *Height* (*Width*) raises the price by 0.091 (0.112) units. Similar for the other variables.
- For each variable the t-test is relative to the null-hypothesis that the true parameter (population parameter) equals 0.
- We fail to reject the null-hypothesis only for *House*.

## Exercise 4

Interpretation of the model:

$$\begin{aligned} Price = -5.520 + 0.091 Height + 0.112 Width \\ +2.292 Signed + 0.389 House + u \end{aligned} \tag{2}$$

- *Ceteris paribus*, 1 inch increase in *Height* (*Width*) raises the price by 0.091 (0.112) units. Similar for the other variables.
- For each variable the t-test is relative to the null-hypothesis that the true parameter (population parameter) equals 0.
- We fail to reject the null-hypothesis only for *House*.
- The theory is not rejected, although our null-hypotheses are $H_0 : \beta_i = 0$.

## Exercise 4

Interpretation of the model:

$$\begin{aligned} Price = & -5.520 + 0.091 Height + 0.112 Width \\ & +2.292 Signed + 0.389 House + u \end{aligned} \tag{2}$$

- *Ceteris paribus*, 1 inch increase in *Height* (*Width*) raises the price by 0.091 (0.112) units. Similar for the other variables.
- For each variable the t-test is relative to the null-hypothesis that the true parameter (population parameter) equals 0.
- We fail to reject the null-hypothesis only for *House*.
- The theory is not rejected, although our null-hypotheses are $H_0 : \beta_i = 0$. One-tailed test would be more appropriate (see script).

## Exercise 4

Do you see any problem in your model?

# Exercise 4

Do you see any problem in your model?

Check the following. In R:

```
qplot(x = model$fitted.values, y = model$residuals) +
  geom_point() + xlab("fitted values") +
  ylab("residuals") +
  geom_hline(yintercept = 0, color = c("red"))
```

# Exercise 4

Do you see any problem in your model?

Check the following. In R:

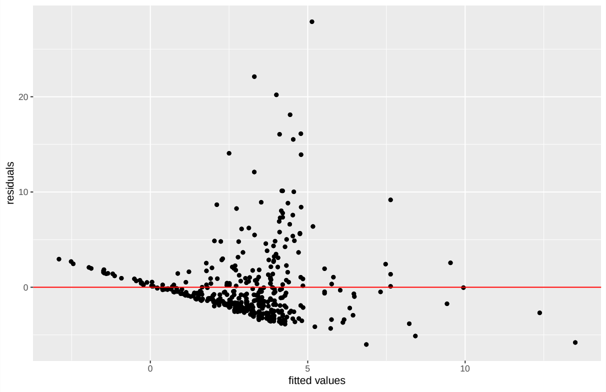
```
qplot(x = model$fitted.values, y = model$residuals) +
  geom_point() + xlab("fitted values") +
  ylab("residuals") +
  geom_hline(yintercept = 0, color = c("red"))
```

In Stata:

```
rvfplot, yline(0)
```

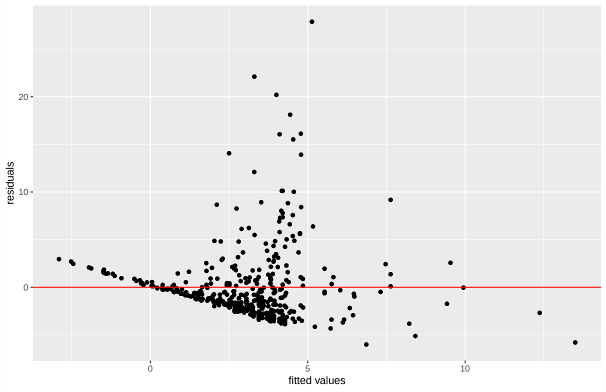
# Exercise 4

What does the following picture suggest to you?



# Exercise 4

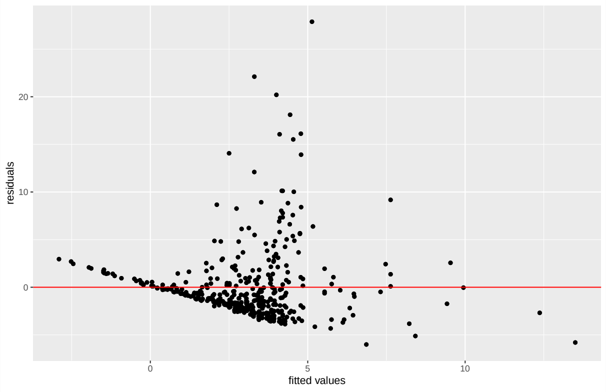
What does the following picture suggest to you?




It is very likely that our *errors* (**not** residuals) are heteroskedastic.

# Exercise 4

What does the following picture suggest to you?




It is very likely that our *errors* (**not** residuals) are heteroskedastic.

We cannot trust the standard errors of our model.

## Exercise 4

Procedure to have heteroskedasticity-robust estimators for the standard errors (and trustworthy t-tests):

# Exercise 4

Procedure to have heteroskedasticity-robust estimators for the standard errors (and trustworthy t-tests):

In R:

```
library(sandwich)
library(lmtest)
robust <- vcovHC(model, type = "HC1")
coeftest(model, vcov. = robust)
```

# Exercise 4

Procedure to have heteroskedasticity-robust estimators for the standard errors (and trustworthy t-tests):

In R:

```
library(sandwich)
library(lmtest)
robust <- vcovHC(model, type = "HC1")
coeftest(model, vcov. = robust)
```

In Stata:

```
reg price height width signed house, robust
ereturn list r2_a
```

# Exercise 4

Heteroskedasticity-robust estimators for the standard errors:

| | *Dependent variable:* | |
|---|---|---|
| | PRICE | |
| | non-robust | robust |
| | (1) | (2) |
| HEIGHT | 0.091*** | 0.091*** |
| | (0.022) | (0.022) |
| WIDTH | 0.112*** | 0.112*** |
| | (0.021) | (0.019) |
| SIGNED | 2.292*** | 2.292*** |
| | (0.503) | (0.349) |
| HOUSE | 0.389 | 0.389 |
| | (0.328) | (0.289) |
| Constant | −5.520*** | −5.520*** |
| | (1.092) | (0.939) |
| Observations | 430 | 430 |
| Adjusted $R^2$ | 0.179 | 0.179 |
| F Statistic (df = 4; 425) | 24.395*** | 24.395*** |
| *Note:* | *p<0.1; **p<0.05; ***p<0.01 | |

## Exercise 5

5. How good is your model, jointly considered?

## Exercise 5

5. How good is your model, jointly considered?

- An F-test is performed on the model.

## Exercise 5

5. How good is your model, jointly considered?

- An F-test is performed on the model. The F-test compares two nested models: a restricted and an unrestricted one.

## Exercise 5

5. How good is your model, jointly considered?

- An F-test is performed on the model. The F-test compares two nested models: a restricted and an unrestricted one.
- We reject the null-hypothesis that the independent variables, *jointly* considered, are not significant in determining the dependent variable.

# Exercise 5

5. How good is your model, jointly considered?

- An F-test is performed on the model. The F-test compares two nested models: a restricted and an unrestricted one.
- We reject the null-hypothesis that the independent variables, *jointly* considered, are not significant in determining the dependent variable.
- The $R^2$ tells how much of the variance of the observations for the dependent variable is explained by the included regressors. (note: we use the *adjusted* $R^2$)

# Exercise 5

5. How good is your model, jointly considered?

- An F-test is performed on the model. The F-test compares two nested models: a restricted and an unrestricted one.
- We reject the null-hypothesis that the independent variables, *jointly* considered, are not significant in determining the dependent variable.
- The $R^2$ tells how much of the variance of the observations for the dependent variable is explained by the included regressors. (note: we use the *adjusted* $R^2$)
- The F-test is a better statistic to evaluate the model, because it is a test which refers to the population.

## Exercise 5

5. How good is your model, jointly considered?

- An F-test is performed on the model. The F-test compares two nested models: a restricted and an unrestricted one.
- We reject the null-hypothesis that the independent variables, *jointly* considered, are not significant in determining the dependent variable.
- The $R^2$ tells how much of the variance of the observations for the dependent variable is explained by the included regressors. (note: we use the *adjusted* $R^2$)
- The F-test is a better statistic to evaluate the model, because it is a test which refers to the population. The $R^2$ has no population counterpart: it merely refers to the sample at hand. Thus its meaning is limited.

# Exercise 5

How do we compute the $R^2$?

# Exercise 5

How do we compute the $R^2$?

$$R^2 = \frac{ESS}{TSS} = \frac{\sum_{i=1}^{n} (\hat{y}_i - \bar{y})^2}{\sum_{i=1}^{n} (y_i - \bar{y})^2} = 1 - \frac{\sum_{i=1}^{n} \hat{u}_i{}^2}{\sum_{i=1}^{n} (y_i - \bar{y})^2}$$

# Exercise 5

How do we compute the $R^2$?

$$R^2 = \frac{ESS}{TSS} = \frac{\sum_{i=1}^{n} (\hat{y}_i - \bar{y})^2}{\sum_{i=1}^{n} (y_i - \bar{y})^2} = 1 - \frac{\sum_{i=1}^{n} \hat{u}_i{}^2}{\sum_{i=1}^{n} (y_i - \bar{y})^2}$$

How do we perform the $F$ test?

# Exercise 5

How do we compute the $R^2$?

$$R^2 = \frac{ESS}{TSS} = \frac{\sum_{i=1}^{n} (\hat{y}_i - \bar{y})^2}{\sum_{i=1}^{n} (y_i - \bar{y})^2} = 1 - \frac{\sum_{i=1}^{n} \hat{u}_i{}^2}{\sum_{i=1}^{n} (y_i - \bar{y})^2}$$

How do we perform the $F$ test?
Start from an $F$ distribution with $k-1$ and $n-k$ degrees of freedom ($k$ independent variables, intercept included, $n$ observations).

# Exercise 5

How do we compute the $R^2$?

$$R^2 = \frac{ESS}{TSS} = \frac{\sum_{i=1}^{n} (\hat{y}_i - \bar{y})^2}{\sum_{i=1}^{n} (y_i - \bar{y})^2} = 1 - \frac{\sum_{i=1}^{n} \hat{u}_i{}^2}{\sum_{i=1}^{n} (y_i - \bar{y})^2}$$

How do we perform the $F$ test?

Start from an $F$ distribution with $k-1$ and $n-k$ degrees of freedom ($k$ independent variables, intercept included, $n$ observations).

The critical value ($F$ statistic) from such $F$ distribution is:

$$F[k-1, n-k] = \frac{R^2/(k-1)}{(1-R^2)/(n-k)}$$

## Exercise 6

6. Based on your model, predict what would be the price of a painting with arbitrary values for the dependent variables

We can do it with pen and paper, we don't really need a software:

# Exercise 6

6. Based on your model, predict what would be the price of a painting with arbitrary values for the dependent variables

We can do it with pen and paper, we don't really need a software:
If $Height = 30$, $Width = 30$, $Signed = 0$, $House = 3$ then:

$$\begin{aligned}\hat{Price} = -5.520 + 0.091 Height + 0.112 Width \\ +2.292 Signed + 0.389 House\end{aligned} \tag{3}$$

$$\begin{aligned}\hat{Price} = -5.520 + 0.091 \times 30 + 0.112 \times 30 \\ +2.292 \times 0 + 0.389 \times 3 = 1.737\end{aligned} \tag{4}$$

# Conclusion

All clear? Questions?

Thanks and see you next week!

## References


Greene, W. H. (2003). *Econometric analysis*. Pearson Education India.

Hlavac, M. (2018). *stargazer: Well-Formatted Regression and Summary Statistics Tables*. Central European Labour Studies Institute (CELSI), Bratislava, Slovakia. R package version 5.2.2.