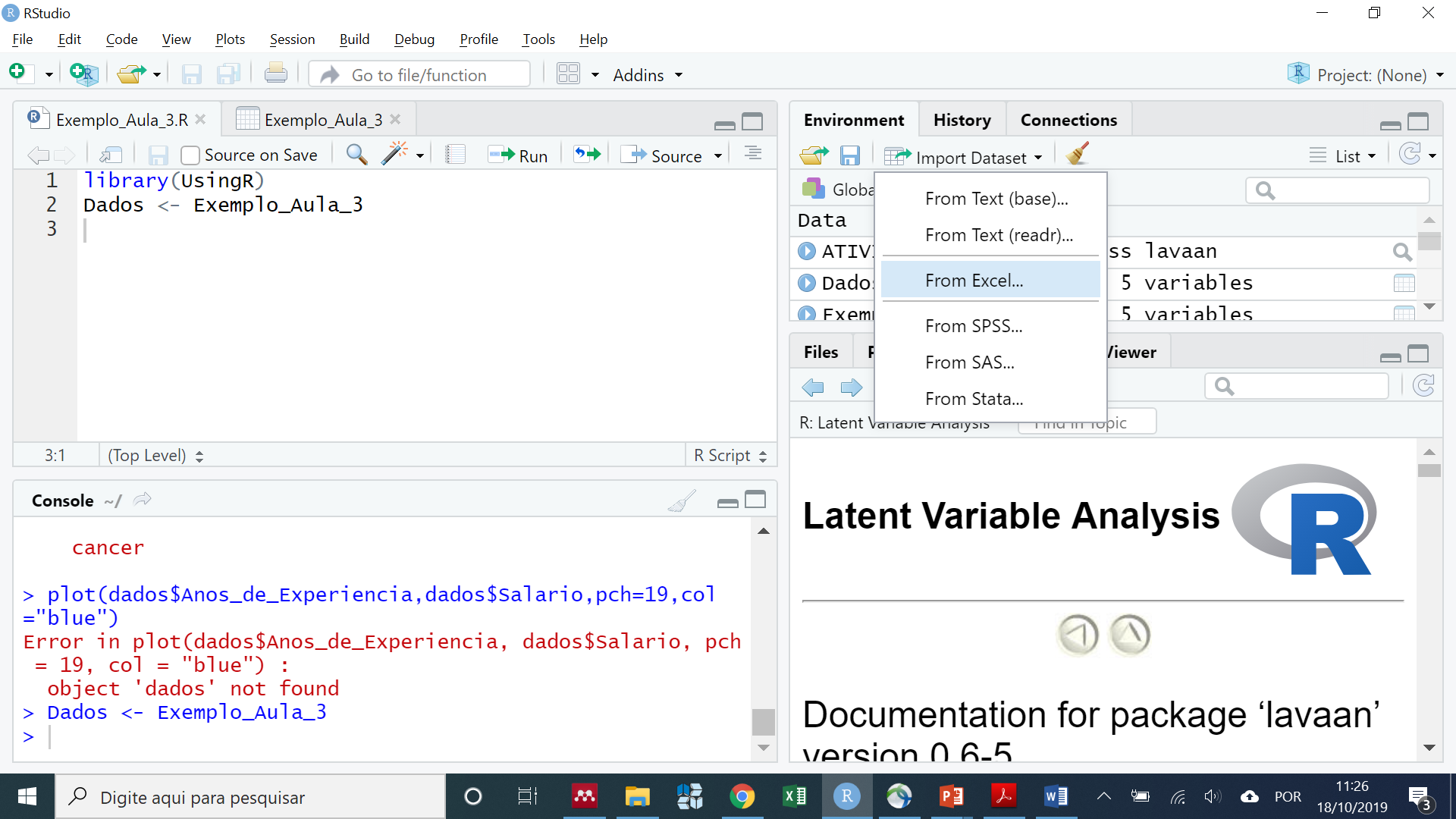
**Utilizando o R para fazer Regressão**

Parte 1 – Regressão Linear Simples

**Carregar o arquivo**



**Instalar pacote com a Regressão**

Para fazer a regressão, precisa instalar um pacote:

install.packages("UsingR")

Carregar o pacote:

library (UsingR)

**Renomear a base de dados**

dados <- Exemplo\_Aula\_3

**Algumas funções de exploração da relação entre as variáveis**

Para entender como as variáveis se relacionam, é possível desenvolver gráficos, como a seguir.

*plot(dados$Anos\_de\_Experiencia,dados$Salario,pch=19,col="blue")*

*plot(dados$Anos\_de\_Experiencia,dados$Salario,pch=20,col="red")*

Obs: pode ser usada a tabela de cores do html

**Colocando a linha de regressão no gráfico**

*plot(dados$Anos\_de\_Experiencia, dados$Salario,*

*xlab = "Anos de Experiência",*

*ylab = "Salário",*

*bg = "red", col = "green", cex = 1.1, pch = 21, frame = FALSE)*

*abline(lm(Salario ~ Anos\_de\_Experiencia, data = dados), lwd = 3)*

observar que:

* bg – cor dos pontos
* col – cor do envoltório dos pontos
* cex – tamanho dos pontos
* pch – tipo dos pontos
* frame = FALSE significa que os eixos não vão se cruzar (fica aberto)
* lwd – espessura da reta de regressão



**Cálculo da Correlação**

*cor(dados$Anos\_de\_Experiencia,dados$Salario)*

[1] 0.8553203

**Calculando os coeficientes de Regressão**

*equation <- lm(dados$Salario ~ dados$Anos\_de\_Experiencia)*

*coef(equation)*

(Intercept) dados$Anos\_de\_Experiencia

22.811124 1.619976

Função lm – primeiro a variável dependente (y) e depois a variável independente(x); ou ainda, primeiro a variável resposta e depois a preditora.

**Para visualizar resumo das informações da análise de regressão:**

*summary(equation)*

Call:

lm(formula = dados$Salario ~ dados$Anos\_de\_Experiencia)

Residuals:

Min 1Q Median 3Q Max

-5.291 -1.441 0.249 0.719 8.849

Coefficients:

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

(Intercept) 22.8111 1.3761 16.576 2.39e-12 \*\*\*

dados$Anos\_de\_Experiencia 1.6200 0.2313 7.004 1.54e-06 \*\*\*

---

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

Residual standard error: 2.991 on 18 degrees of freedom

Multiple R-squared: 0.7316, Adjusted R-squared: 0.7167

F-statistic: 49.06 on 1 and 18 DF, p-value: 1.541e-06