Contenido

[1. Obtenga el mejor modelo de regresión lineal simple basado en las variables meteorológicas. 3](#_Toc173318793)

[Modelos para las distintas variables meteorológicas: 4](#_Toc173318794)

[Validación de los supuestos para el **“Modelo3”** 6](#_Toc173318795)

[Explicación del Modelo3 9](#_Toc173318796)

[Test para ro usando cor.test 10](#_Toc173318797)

[2. Obtenga el mejor modelo de regresión lineal simple basado en los contaminantes atmosféricos. 12](#_Toc173318798)

[Modelos para los distintos contaminantes atmosféricos: 12](#_Toc173318799)

[Validación de los supuestos para el **“Modelo2”** 15](#_Toc173318800)

[Explicación del Modelo2 17](#_Toc173318801)

[Test para ro usando cor.test 18](#_Toc173318802)

[3. Con base a todas las variables (meteorológicas y contaminantes), mediante una técnica iterativa (forward o backward) seleccione el mejor modelo predictivo. Indique para cada paso qué variable entra/sale del modelo, indicando el aumento/disminución del R2-ajustado. 20](#_Toc173318803)

[Resumen con comando summary para el M2 20](#_Toc173318804)

[Test con anova comparando modelo **completo** vs **M2** 21](#_Toc173318805)

[Detalle del R2 Ajustado por paso a paso del step 22](#_Toc173318806)

[4. Basado en los resultados previos, proponga un modelo con tres predictores (debe incluir una variable meteorológica y dos contaminantes), revise supuestos y evalúe con especial énfasis el problema de multicolinealidad. Apóyese de tablas de correlación, gráficos y métricas respectivas. 23](#_Toc173318807)

[Resumen de los R2 para variables meteorológica 23](#_Toc173318808)

[Resumen de los R2 para variables contaminantes 23](#_Toc173318809)

[**Modelo con 3 predictores:** 24](#_Toc173318810)

[Validación de los supuestos para el ModeloEscogido PM2.5 ~ TMin + NO2 + CO 27](#_Toc173318811)

[Análisis de multicolinealidad 29](#_Toc173318812)

[Cálculo de Correlaciones 31](#_Toc173318813)

[Matriz de correlación 31](#_Toc173318814)

[Coeficiente de determinación del modelo 31](#_Toc173318815)

[Pruebas de Hipótesis para Correlaciones (𝜌) 32](#_Toc173318816)

[Valores atípicos o influyentes 33](#_Toc173318817)

[Cook 34](#_Toc173318818)

[Leverage 34](#_Toc173318819)

**Alumno: Rodrigo Jeldres Carrasco**

Control N3

Variable respuesta PM2.5

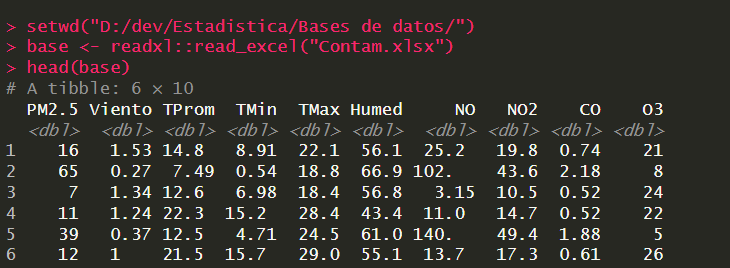
El objetivo es entender y explicar el comportamiento de los niveles de contaminación del aire en la Región Metropolitana. Para ello, desde el Sistema de Información Nacional de Calidad del Aire (sinca.mma.gob.cl), se seleccionó una muestra de la información histórica de la Estación Parque O’Higgins en Santiago, la cual ha sido almacenada en la base ***Contam.xlsx.*** Se dispone de las siguientes variables:

* **PM2.5** – Materia particulada de 2.5 mg/m3. El PM2.5 son partículas muy pequeñas suspendidas en el aire que tienen un diámetro de menos de 2.5 micras. La materia particulada incluye sustancias químicas orgánicas, polvo, hollín y metales. Es nuestra variable respuesta.
* Potenciales variables explicativas:

|  |  |
| --- | --- |
| **Variables meteorológicas** | **Contaminantes atmosféricos** |
| Viento – Velocidad del viento (m/s) | NO – Monóxido de nitrógeno (ppb) |
| TProm – Temperatura promedio (° Celsius) | NO2 – Dióxido de nitrógeno (ppb) |
| TMin – Temperatura mínima (° Celsius) | CO – Monóxido de carbono (ppm) |
| TMax – Temperatura máxima (° Celsius) | O3 – Ozono (ppb) |
| Humed – Humedad relativa del aire (%) |  |

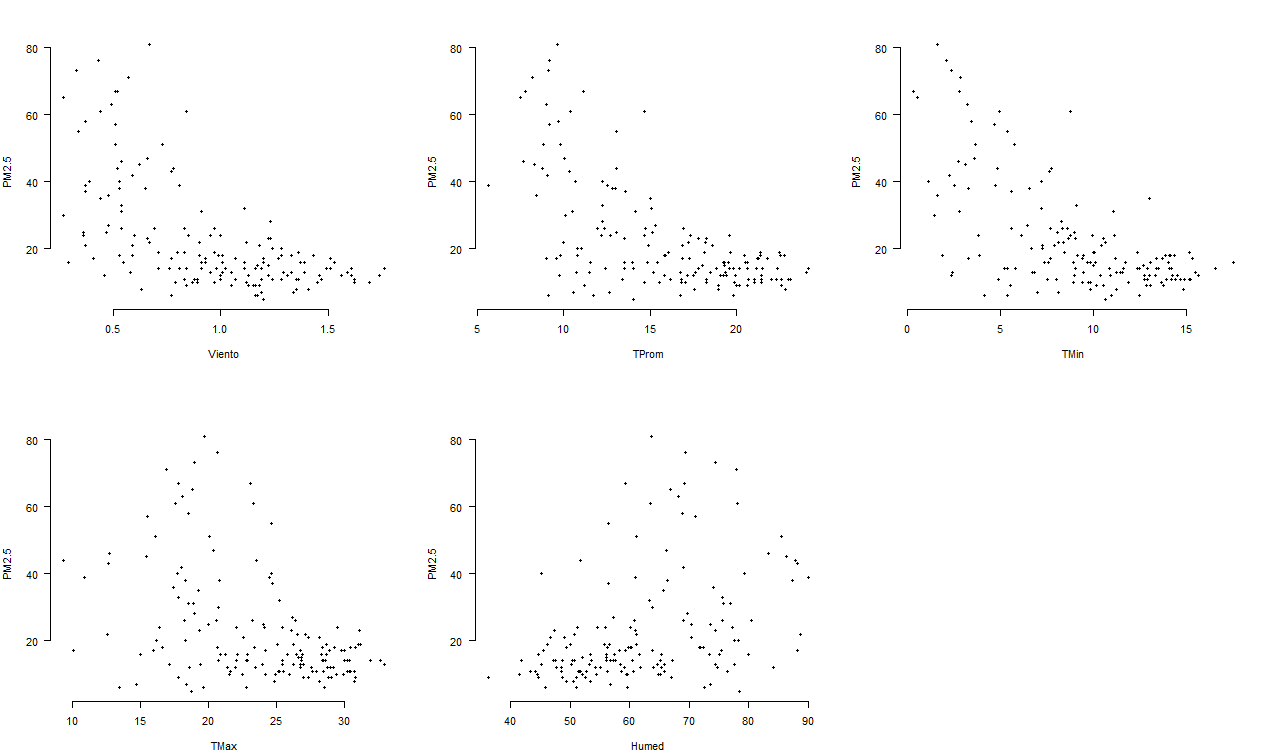
1. Obtenga el mejor modelo de regresión lineal simple basado en las variables meteorológicas.

Primero generar toda la selección de variable a explicar PM2.5 contra las **variables meteorológicas**



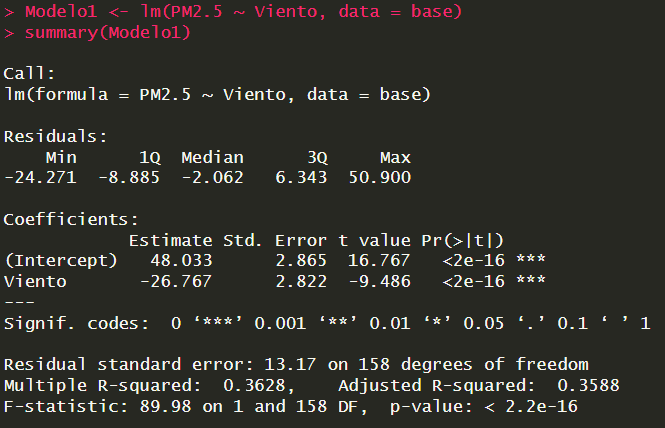
En este caso es Viento, TProm, TMin, TMax, Humed

Gráficamente:



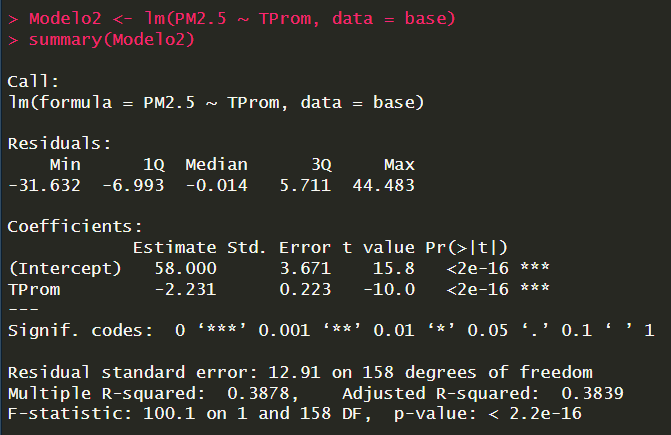
## Modelos para las distintas variables meteorológicas:

1. **Viento**



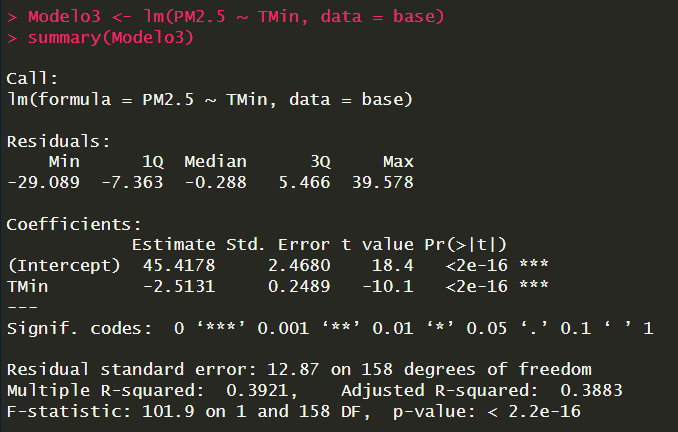
Multiple R-squared: 0.3628

1. **TProm**



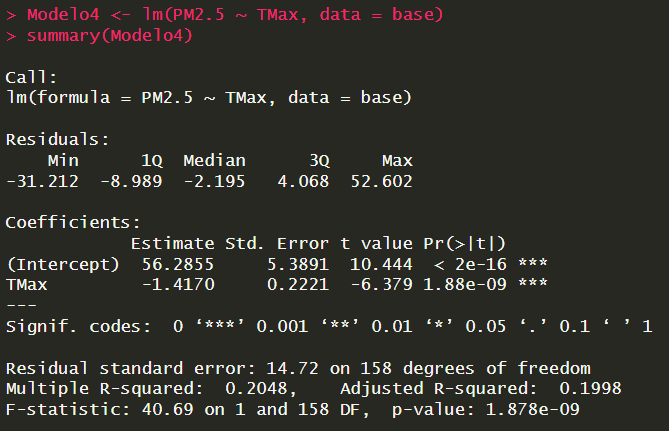
Multiple R-squared: 0.3878

1. **TMin**



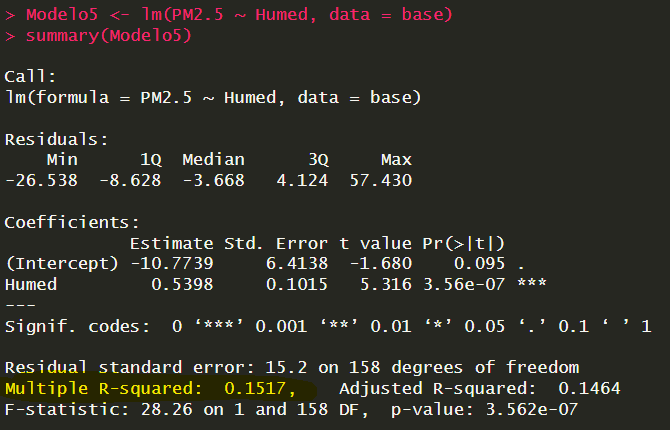
Multiple R-squared: 0.3921

1. **TMax**



Multiple R-squared: 0.2048

1. **Humed**



Multiple R-squared: 0.1517

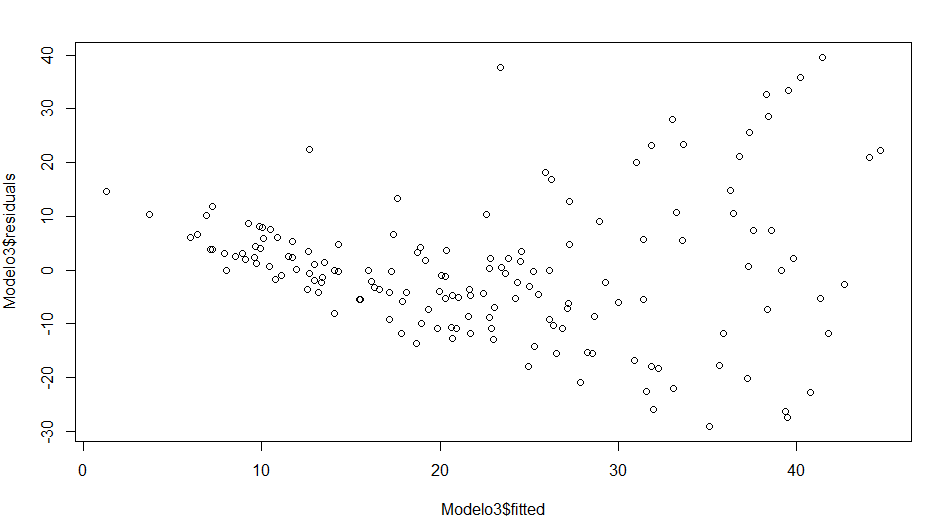
Resp: el mejor modelo simple sin transformación con variable meteorológica es presentado con Modelo3 usando “**Tmin**” dado que tiene la mayor variable explicativa que es 39,21% (en R da con decimales: 0.3921017)

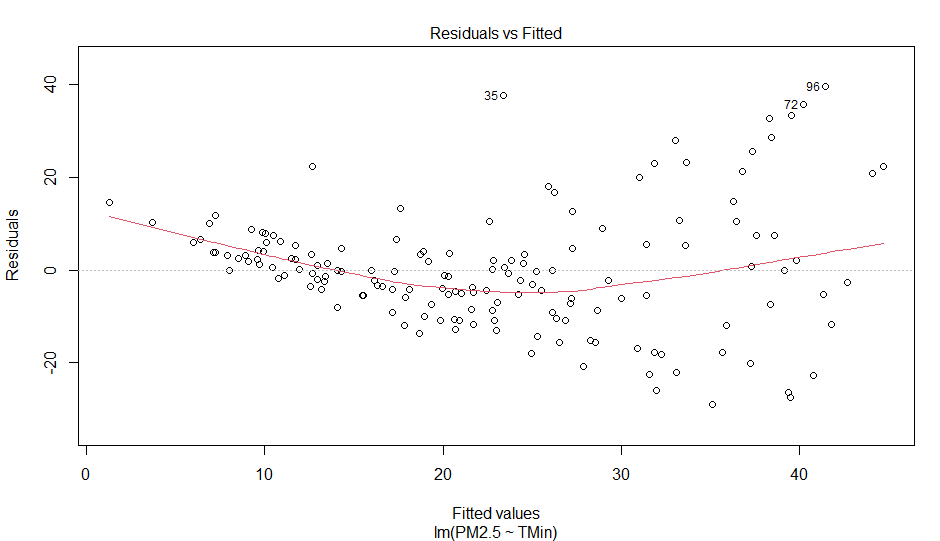
Modelo3 <- lm(PM2.5 ~ TMin, data = base)

## Validación de los supuestos para el **“Modelo3”**

1. **Linealidad**

**plot(Modelo3$fitted, Modelo3$residuals)**



**plot(Modelo3, 1)**

Resp: Cumple linealidad

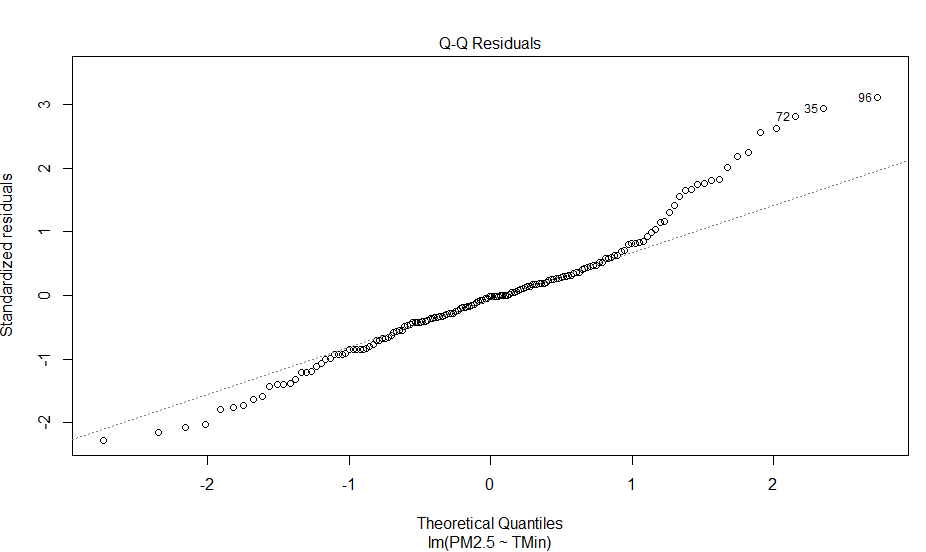
1. **Normalidad**

H0: residuos distribuyen normales

H1: residuos no distribuyen normales

Gráficamente:

**plot(Modelo3, 2)**



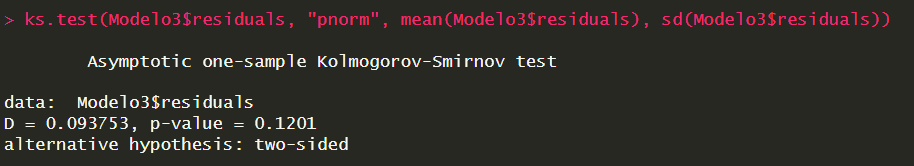
**> nortest::lillie.test(Modelo3$residuals)**

data: Modelo3$residuals

D = 0.093753, p-value = 0.001558

Para un alfa < 0.05, se rechaza H0, es decir, no existe normalidad en los residuos según lillie.test

Test Ks



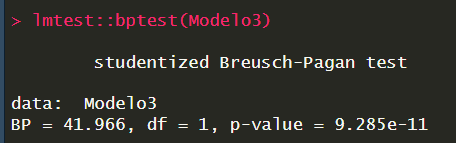
D = 0.093753, p-value = 0.1201, SÍ existe normalidad según Test KS

Resp: no hay normalidad en los residuos

1. **Homocedasticidad**

H0: Sí existe Homocedasticidad

H1: No existe Homocedasticidad

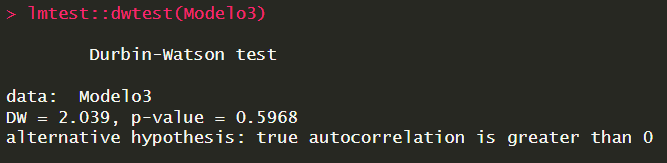


BP = 41.966, df = 1, p-value = 9.285e-11 < alfa = 0.05

Resp: Se rechaza H0, no existe homocedasticidad

1. **Independencia**

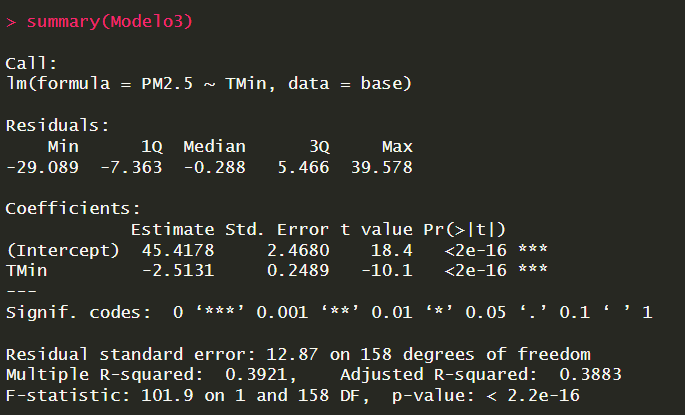
H0: no hay autocorrelación vs H1: hay autocorrelación



DW = 2.039, p-value = 0.5968 > alfa = 0.05

Resp: se acepta H0, existe independencia

## Explicación del Modelo3



Para el Beta1: B1 = - 2.5131

Modelo3 quedaría con:

**Y = B0 + B1 \* X**

**PM2.5 = 45,42 – 2,51 \* TMin**

**Test de significancia del B1**

H0: **B1 = 0**

H1: **B1 distinto de 0**

Salida R:

t-value = -10.1 // p-value: <2e-16 < alfa = 0.05, se rechaza H0

Para un alfa tan pequeño, la variable **TMin** es significativa

**Test de significancia del modelo**

H0: no existe regresión vs H1: existe regresión

F-statistic: 101.9 on 1 and 158 DF, p-value: < 2.2e-16

< alfa = 0.05, se rechaza H0, es decir, existe regresión

**Multiple R-squared: 0.3921**: el modelo explica en un 39,21% el nivel de contaminación PM2.5

## Test para ro usando cor.test

**H0: ro = 0**

**H1: ro distinto de 0**

**# Correlación entre PM2.5 y Viento**

cor.test(base$PM2.5, base$Viento)

# t = -9.4857, p-value < 2.2e-16

**# Correlación entre PM2.5 y TProm**

cor.test(base$PM2.5, base$TProm)

# t = -10.004, p-value < 2.2e-16

**# Correlación entre PM2.5 y TMin**

cor.test(base$PM2.5, base$TMin)

# t = -10.095, p-value < 2.2e-16

**# Correlación entre PM2.5 y TMax**

cor.test(base$PM2.5, base$TMax)

# t = -6.3791, p-value = 1.878e-09

**# Correlación entre PM2.5 y Humed**

cor.test(base$PM2.5, base$Humed)

# t = 5.3161, p-value = 3.562e-07

**Interpretación:**

PM2.5 y Viento: Correlación negativa significativa (t = -9.4857, p-value < 2.2e-16).

PM2.5 y TProm: Correlación negativa significativa (t = -10.004, p-value < 2.2e-16).

PM2.5 y TMin: Correlación negativa significativa (t = -10.095, p-value < 2.2e-16).

PM2.5 y TMax: Correlación negativa significativa (t = -6.3791, p-value = 1.878e-09).

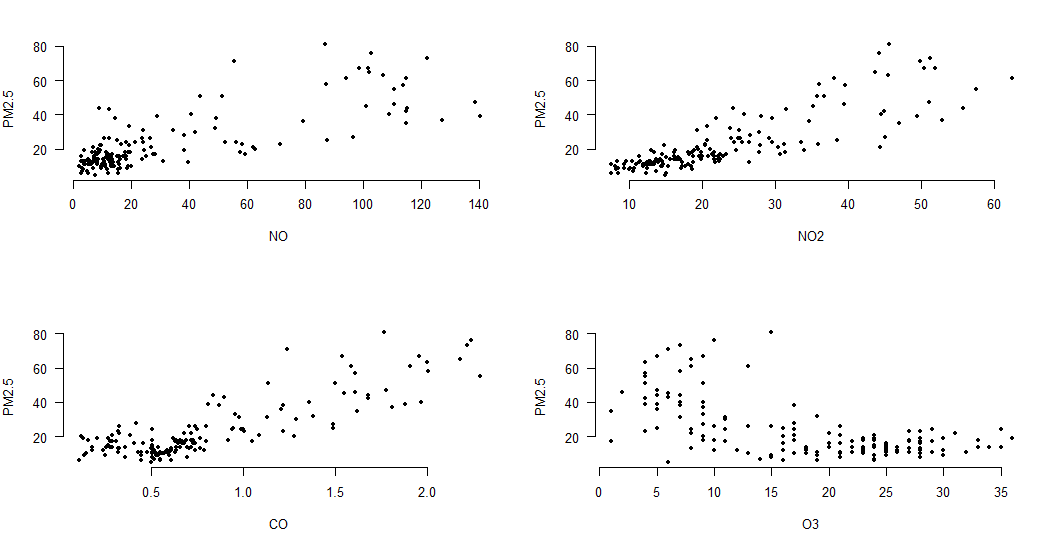
PM2.5 y Humed: Correlación positiva significativa (t = 5.3161, p-value = 3.562e-07).

Cada uno de estos tests sugiere que hay una correlación significativa entre las variables analizadas y **PM2.5**, dado que todos los p-values son mucho menores que el nivel de significancia típico de 0.05. En todos los casos, H0 es rechazado.

1. Obtenga el mejor modelo de regresión lineal simple basado en los contaminantes atmosféricos.

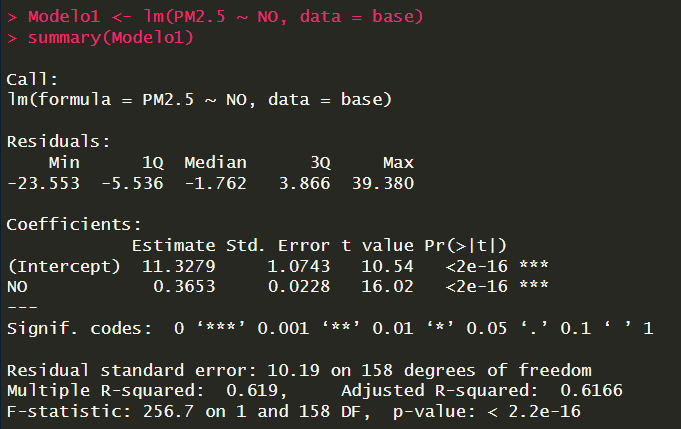
|  |
| --- |
| **Contaminantes atmosféricos** |
| NO – Monóxido de nitrógeno (ppb) |
| NO2 – Dióxido de nitrógeno (ppb) |
| CO – Monóxido de carbono (ppm) |
| O3 – Ozono (ppb) |

Gráficamente:



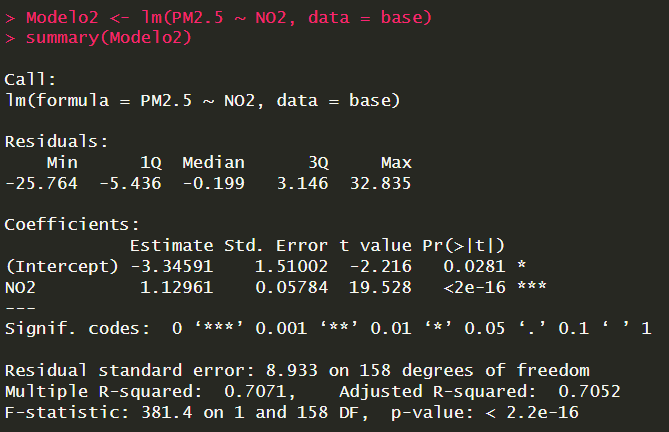
## Modelos para los distintos contaminantes atmosféricos:

1. **NO**



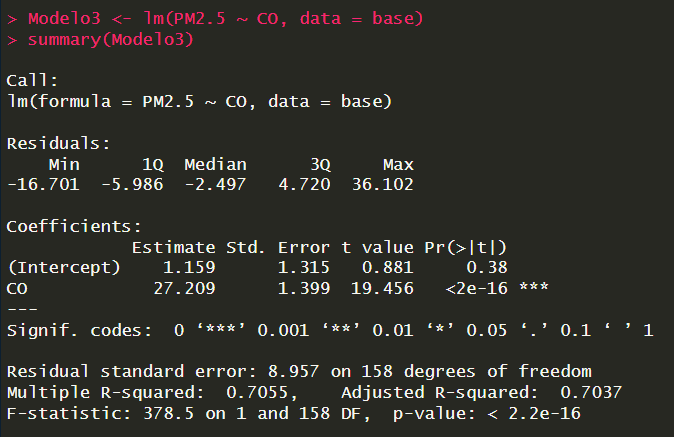
Multiple R-squared: 0.619

1. **NO2**



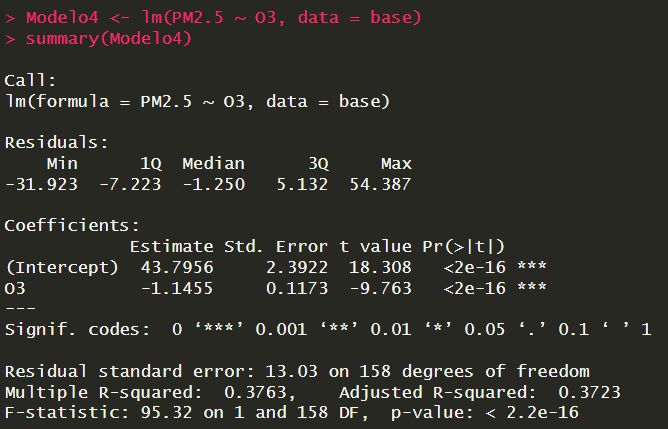
Multiple R-squared: 0.7071

1. **CO**



Multiple R-squared: 0.7055

1. **O3**



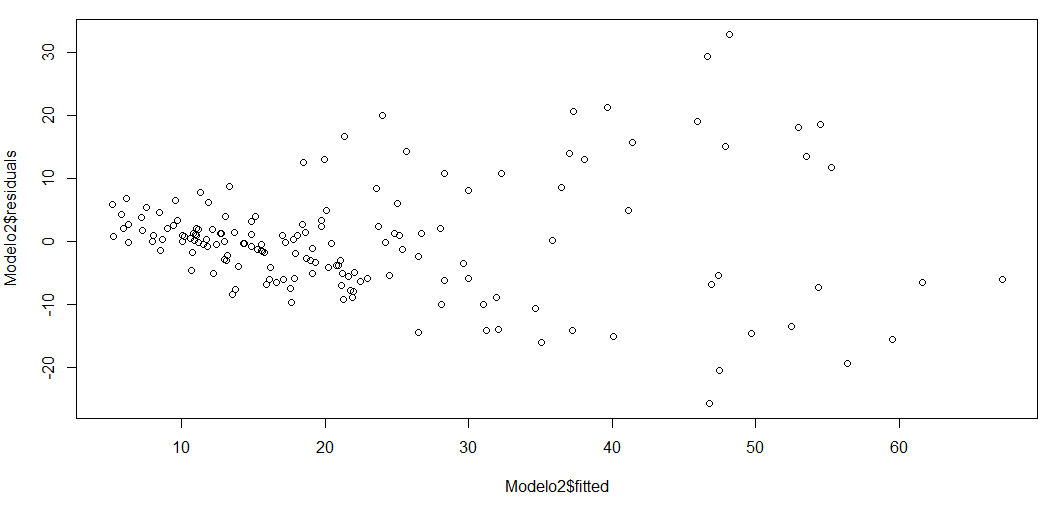
Multiple R-squared: 0.3763

Resp: el mejor modelo simple sin transformación con variable contaminantes atmosféricos es presentado con Modelo2 usando “NO2” dado que tiene la mayor R2 que es 70,7% (en R da con decimales: 0.7070606)

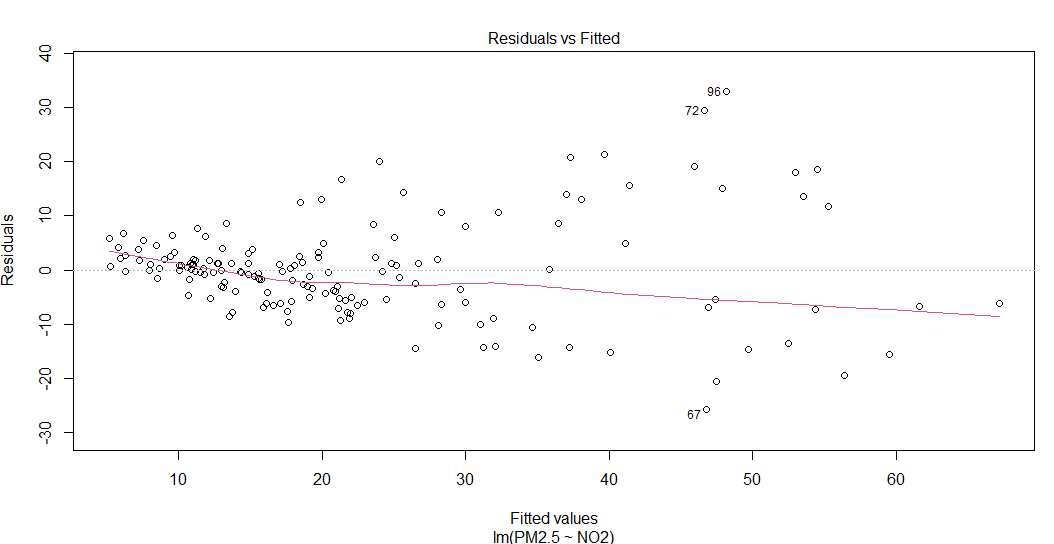
## Validación de los supuestos para el **“Modelo2”**

1. **Linealidad**

**plot(Modelo2$fitted, Modelo2$residuals)**



**plot(Modelo2, 1)**



Resp: Se acepta la linealidad

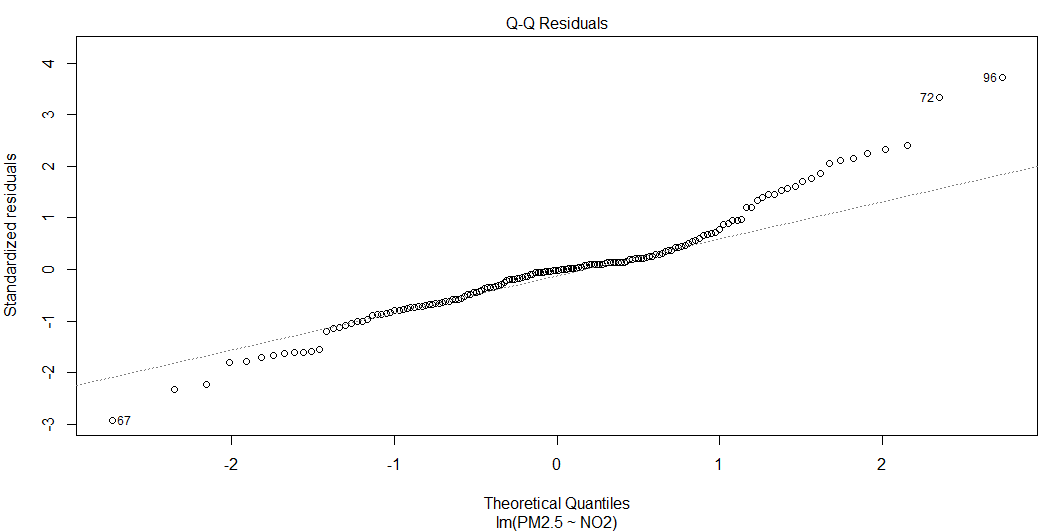
1. **Normalidad**

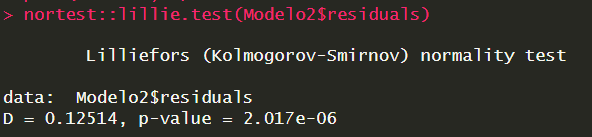
H0: residuos distribuyen normales

H1: residuos no distribuyen normales

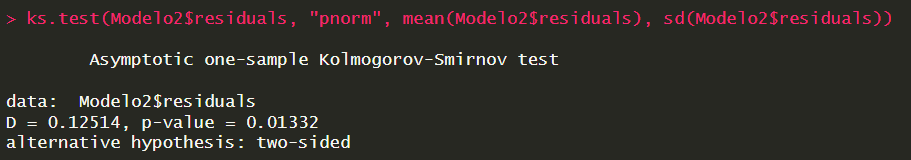
Gráficamente:

**plot(Modelo2, 2)**





D = 0.12514, p-value = 2.017e-06 < alfa = 0.05 Se rechaza H0



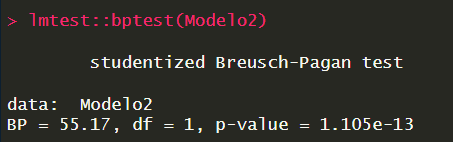
D = 0.12514, p-value = 0.01332 < alfa = 0.05 se rechaza H0

Resp: En ambos test, se rechaza H0, es decir, residuos no distribuyen normales

1. **Homocedasticidad**

H0: Sí existe Homocedasticidad

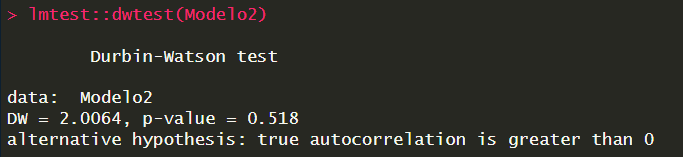
H1: No existe Homocedasticidad



BP = 55.17, df = 1 ; p-value = 1.105e-13 < alfa = 0.05, se rechaza H0, es decir, No existe Homocedasticidad

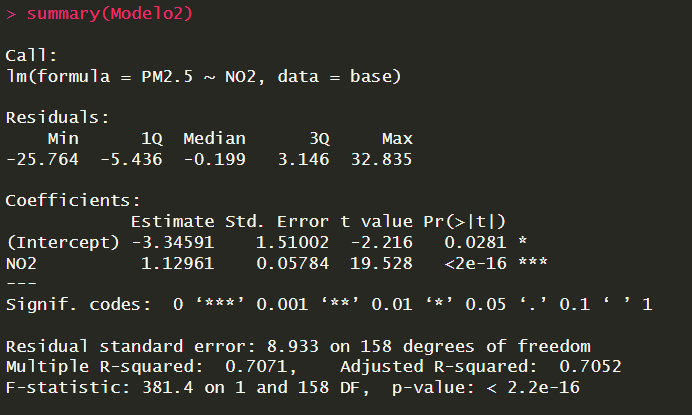
1. **Independencia**

H0: no hay autocorrelación vs H1: hay autocorrelación



DW = 2.0064, p-value = 0.518 > alfa = 0.05 Se acepta H0, existe independencia

## Explicación del Modelo2



B1 = 1.12961

Modelo2 quedaría con:

**Y = B0 + B1 \* X**

**PM2.5 = -3.35 + 1.13 \* NO2**

**Test de significancia del B1**

H0: **B1 = 0**

H1: **B1 distinto de 0**

Salida R:

(NO2) t-value= 19,528; P-VALUE= < 2e-16 es menor a alfa = 0,05; Se rechaza H0

Para un alfa tan pequeño, la variable **NO2** es significativa

**Test de significancia del modelo**

H0: no existe regresión vs H1: existe regresión

F-statistic: 381,4; p-value: < 2.2e-16 ; es menor a alfa = 0,05, se rechaza H0 , es decir , existe regresión

**Multiple R-squared**: 0,7071: el modelo explica en un 70,71% el nivel de contaminación PM2.5

## Test para ro usando cor.test

**H0: ro = 0**

**H1: ro distinto de 0**

# Correlación entre PM2.5 y NO

cor.test(base$PM2.5, base$NO)

# t = 16;021; p-value < 2.2e-16

# Correlación entre PM2.5 y NO2

cor.test(base$PM2.5, base$NO2)

# t = 19,528; p-value < 2.2e-16

# Correlación entre PM2.5 y CO

cor.test(base$PM2.5, base$CO)

# t = 19,456; p-value < 2.2e-16

# Correlación entre PM2.5 y O3

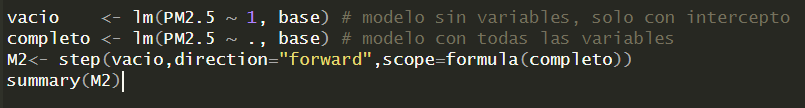
cor.test(base$PM2.5, base$O3)

# t = -9,7629; p-value < 2.2e-16

**Interpretación**: Dado un nivel de significancia α = 0.05, todas las correlaciones entre las concentraciones de PM2.5 y las variables NO, NO2, CO y O3 son estadísticamente significativas. Esto se confirma por los p-values menores a 0.05 en cada caso, lo que indica que podemos rechazar la hipótesis nula de que no existe correlación entre PM2.5 y cada una de estas variables.

1. Con base a todas las variables (meteorológicas y contaminantes), mediante una técnica iterativa (forward o backward) seleccione el mejor modelo predictivo. Indique para cada paso qué variable entra/sale del modelo, indicando el aumento/disminución del R2-ajustado.

**En este caso se usa “forward”**



**Step 1: PM2.5 ~ NO2**: Se agrega NO2

**Step 2: PM2.5 ~ NO2 + CO**: Se agrega CO

**Step 3: PM2.5 ~ NO2 + CO + Humed**: Se agrega Humed

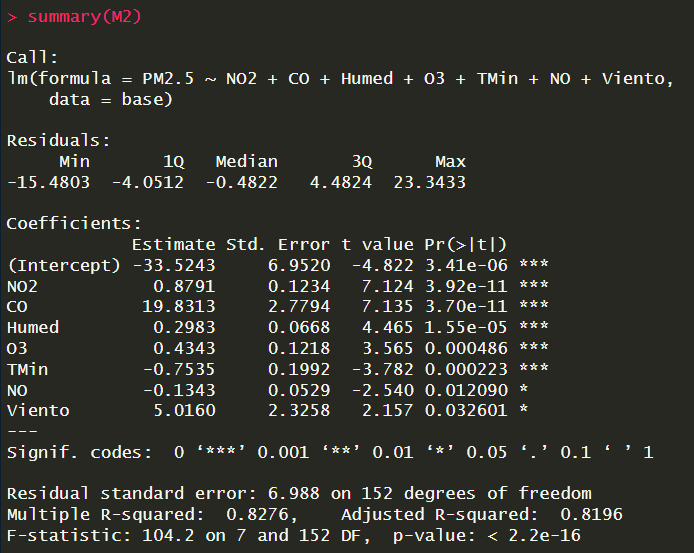
**Step 4: PM2.5 ~ NO2 + CO + Humed + O3**: Se agrega O3

**Step 5: PM2.5 ~ NO2 + CO + Humed + O3 + TMin**: Se agrega TMin

**Step 6: PM2.5 ~ NO2 + CO + Humed + O3 + TMin + NO**: Se agrega NO

**Step 7: PM2.5 ~ NO2 + CO + Humed + O3 + TMin + NO + Viento**: Se agrega Viento

## Resumen con comando summary para el M2

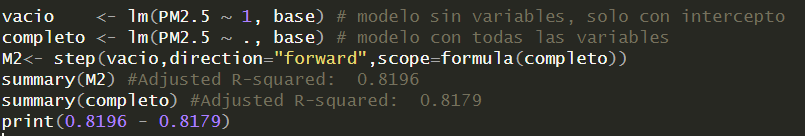


Adjusted R-squared: 0,8196 = 81.96% de calidad del modelo

Calculo del aumento del R2-ajustado

ΔR2-Ajustado=R2(M2)−R2(Completo)

ΔR-Ajustado​ = 0.8196 − 0.8179=0.0017

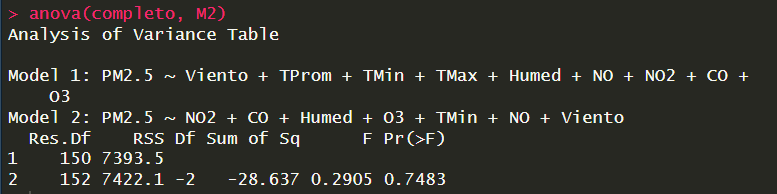


Resp: Para el caso “forward” hubo un aumento del R2 Ajustado de 0.0017 (0.17%). El modelo **M2**, que es más simple y se ha obtenido mediante una selección “forward”, proporciona un ajuste ligeramente mejor a los datos que el modelo **completo**.

## Test con anova comparando modelo **completo** vs **M2**

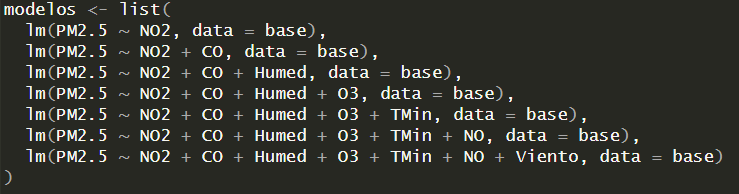
**H₀**: El modelo completo no proporciona un ajuste significativamente mejor que el modelo reducido.

**H₁**: El modelo completo proporciona un ajuste significativamente mejor que el modelo reducido.



**Conclusión:** T = 0.29; p-valor = 0.7483: Esto es mayor que alfa = 0.05, por lo que no se rechaza la hipótesis nula. Esto significa que el modelo **completo** no mejora significativamente el ajuste en comparación con el modelo más sencillo (**M2**). Por lo tanto, el modelo **M2** es preferible en términos de simplicidad y ajuste.

## Detalle del R2 Ajustado por paso a paso del step



Step 1: 0.7052066

Step 2: 0.7627398

Step 3: 0.7877649

Step 4: 0.8011568

Step 5: 0.8115301

Step 6: 0.8153084

Step 7: 0.8196134

Step final con modelo **PM2.5 ~ NO2 + CO + Humed + O3 + TMin + NO + Viento**

**R2 =** 0.8196134

Modelo para este caso:

**Y^=β0​+β1​NO2+β2​CO+β3​Humed+β4​O3+β5​TMin+β6​NO+β7​Viento**

Y^=​−33.5243+0.8791×NO2+19.8313×CO+0.2983×Humed+0.4343×O3−0.7535×TMin−0.1343×NO+5.0160×Viento​

1. Basado en los resultados previos, proponga un modelo con tres predictores (debe incluir una variable meteorológica y dos contaminantes), revise supuestos y evalúe con especial énfasis el problema de multicolinealidad. Apóyese de tablas de correlación, gráficos y métricas respectivas.

## Resumen de los R2 para variables meteorológica

|  |  |
| --- | --- |
| Modelo | Multiple R-squared |
| Modelo1 <- lm(PM2.5 ~ **Viento**, data = base) | 0.3628462 |
| Modelo2 <- lm(PM2.5 ~ **TProm**, data = base) | 0.3877903 |
| Modelo3 <- lm(PM2.5 ~ **TMin**, data = base) | 0.3921017 |
| Modelo4 <- lm(PM2.5 ~ **TMax**, data = base) | 0.2048031 |
| Modelo5 <- lm(PM2.5 ~ **Humed**, data = base) | 0.1517276 |

## Resumen de los R2 para variables contaminantes

|  |  |
| --- | --- |
| Modelo | Multiple R-squared |
| Modelo1 <- lm(PM2.5 ~ **NO**, data = base) | 0.6189904 |
| Modelo2 <- lm(PM2.5 ~ **NO2**, data = base) | 0.7070606 |
| Modelo3 <- lm(PM2.5 ~ **CO**, data = base) | 0.7055177 |
| Modelo4 <- lm(PM2.5 ~ **O3**, data = base) | 0.3762709 |

Supuestos: como el enunciado pide seleccionar una variable meteorológica y dos contaminantes, se escoge los que tienen mejores R2, es decir, el más alto ya que con eso indica un mayor % de explicación del modelo.

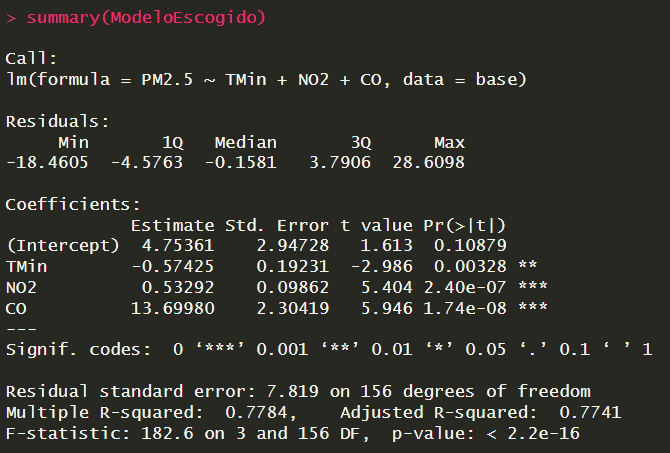
Para el caso de variable meteorológica: **Tmin**

Para el caso de variable contaminante: **NO2** y **CO**

**ModeloEscogido <- lm(PM2.5 ~ TMin + NO2 + CO, data = base)**

## **Modelo con 3 predictores:**

**summary(ModeloEscogido)**



Definición de Ŷ con los 3 predictores

PM2.5 = 4.75 - 0.57 \* TMin + 0.53 \* NO2 + 13.70 \* CO

Adjusted R-squared: 0.7741

Todas las variables son significativas, calidad del modelo ajustado es de 77,41%

**Test de significancia del Bi**

Para cada Bi:

H0: **Bi = 0**

H1: **Bi distinto de 0**

**Coeficiente del Intercepto (constante)**

* Resultado del test: t = 1.613, p-value = 0.10879

Conclusión: No se rechaza la hipótesis nula, ya que el p-valor es mayor que el nivel de significancia de 0.05. No hay suficiente evidencia para concluir que el intercepto es significativamente diferente de cero.

**Coeficiente de TMin**

* Resultado del test: t = -2.986, p-value = 0.00328

Conclusión: Se rechaza la hipótesis nula, ya que el p-valor es menor que el nivel de significancia de 0.05. Hay suficiente evidencia para concluir que el coeficiente de TMin es significativamente diferente de cero.

**Coeficiente de NO2**

* Resultado del test: t = 5.404, p-value = 2.40e-07

Conclusión: Se rechaza la hipótesis nula, ya que el p-valor es mucho menor que 0.05. Hay suficiente evidencia para concluir que el coeficiente de NO2 es significativamente diferente de cero.

**Coeficiente de CO**

* Resultado del test: t = 5.946, p-value = 1.74e-08

Conclusión: Se rechaza la hipótesis nula, ya que el p-valor es mucho menor que 0.05. Hay suficiente evidencia para concluir que el coeficiente de CO es significativamente diferente de cero.

**Conclusión:** Los coeficientes de las variables TMin, NO2 y CO son significativamente diferentes de cero, lo que indica que estas variables tienen un impacto significativo en la variable dependiente PM2.5

**Test de significancia del modelo**

H0: no existe regresión vs H1: existe regresión

F-statistic: 182,6 ; p-value: < 2.2e-16

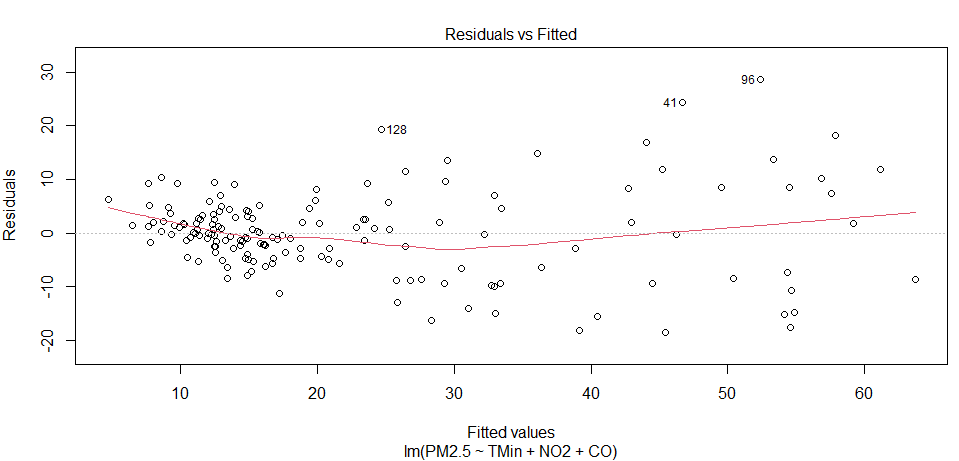
es menor a alfa = 0.05. Se rechaza la hipótesis nula. Esto implica que al menos una de las variables independientes en el modelo tiene un efecto significativo sobre la variable dependiente PM2.5

**Adjusted R-squared:** 0.7741. El modelo con los 3 predictores explica en un 77,41% el nivel de contaminación PM2.5

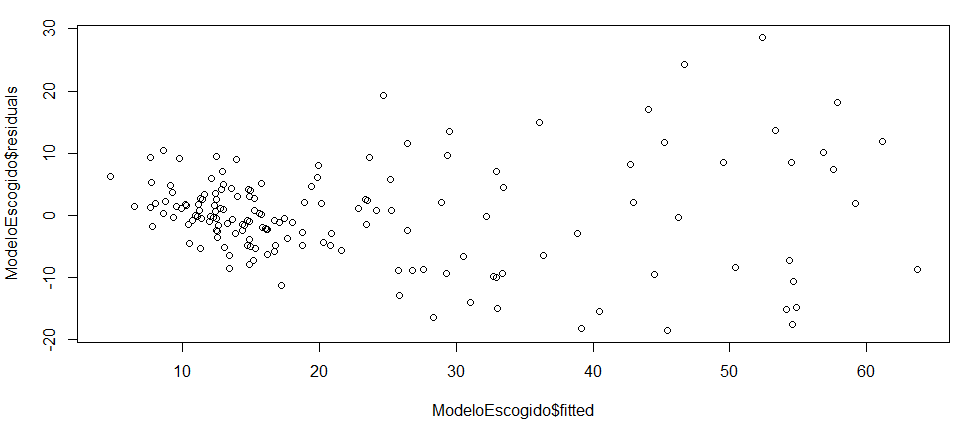
## Validación de los supuestos para el ModeloEscogido PM2.5 ~ TMin + NO2 + CO

1. **Linealidad**

**plot(ModeloEscogido, 1)**



**plot(ModeloEscogido$fitted, ModeloEscogido$residuals)**



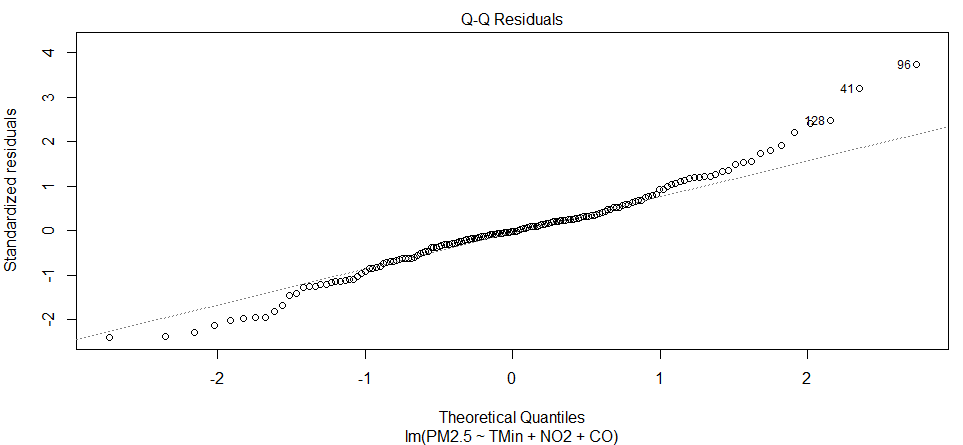
Resp: cumple linealidad

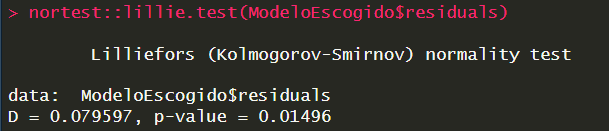
1. **Normalidad**

H0: residuos distribuyen normales

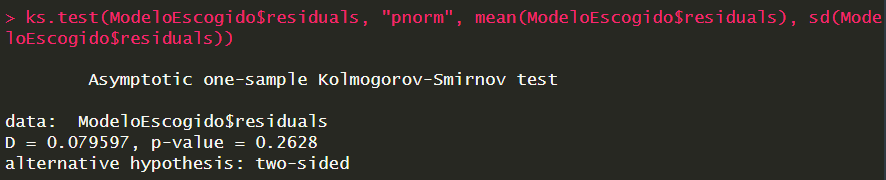
H1: residuos no distribuyen normales

Gráficamente:

**plot(ModeloEscogido, 2)**



D = 0.079597, p-value = 0.01496 < alfa = 0.05, se rechaza H0, no hay normalidad usando lilli test



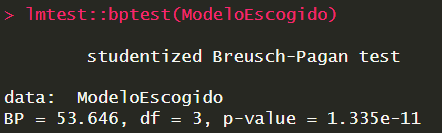
D = 0.079597, p-value = 0.2628 > alfa = 0.05 se acepta H0, hay normalidad según test ks

**Resp:** Los dos tests ofrecen resultados contradictorios respecto a la normalidad de los residuos. El test de Lilliefors indica una falta de normalidad, mientras que el test de KS no encuentra evidencia suficiente para rechazar la normalidad. Se tomará como válido el test Lillie, Se rechaza H0, No existe normalidad de los residuos

1. **Homocedasticidad**

H0: Sí existe Homocedasticidad

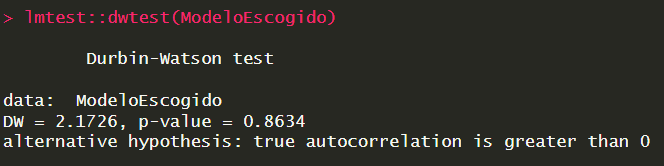
H1: No existe Homocedasticidad



BP = 53.646, df = 3, p-value = 1.335e-11 < alfa = 0.05, se rechaza H0, no existe homocedasticidad

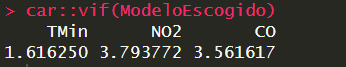
1. **Independencia**

H0: no hay autocorrelación vs H1: hay autocorrelación

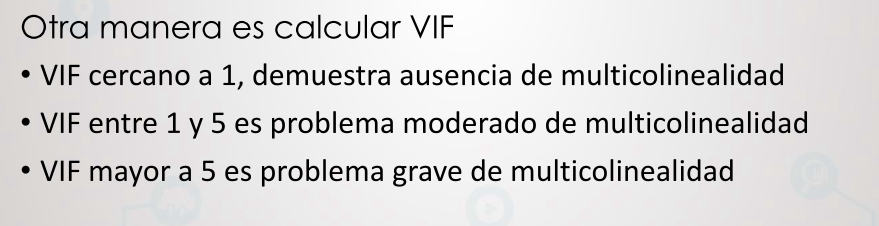


DW = 2.1726, p-value = 0.8634 > alfa = 0.05, Se acepta H0, existe independencia

## Análisis de multicolinealidad



Según documentación



Resp: No hay multicolinealidad grave en el modelo, ya que todos los VIF están por debajo de 5. Dado el rango anterior, cae en la categoría de “moderado”

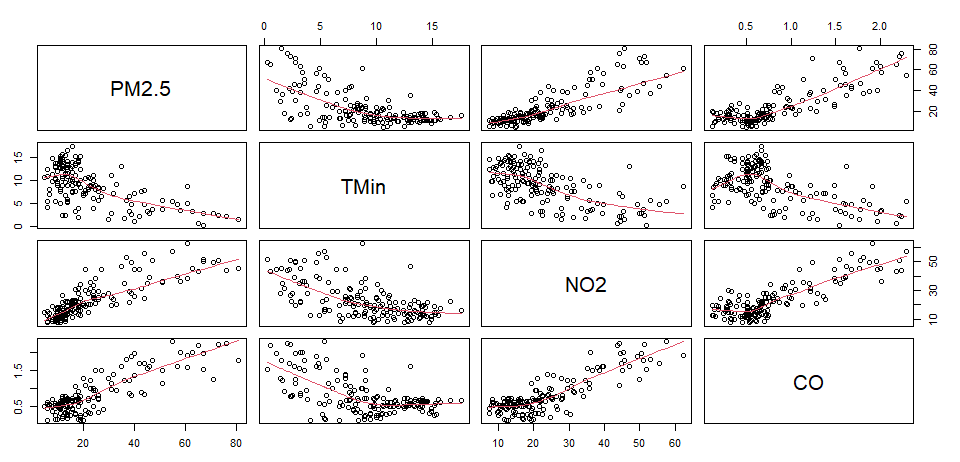
Gráficos de relación entre todas las variables

Análisis de correlación

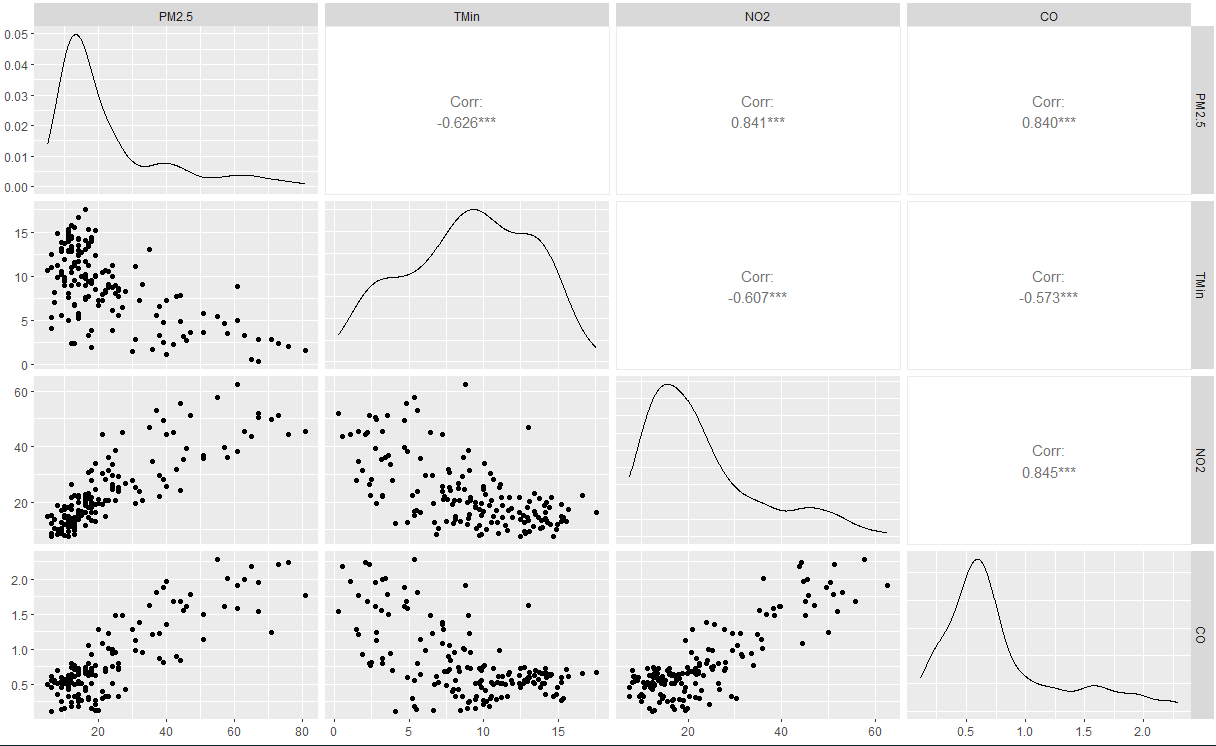
base\_filtrada <- base[, c("PM2.5", "TMin", "NO2", "CO")]

Matriz de dispersión

**pairs(base\_filtrada,upper.panel= panel.smooth, lower.panel = panel.smooth)**



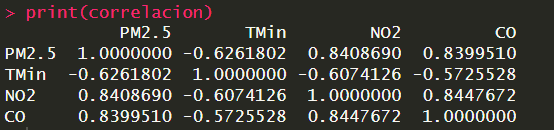
**ggpairs(base\_filtrada)**



## Cálculo de Correlaciones

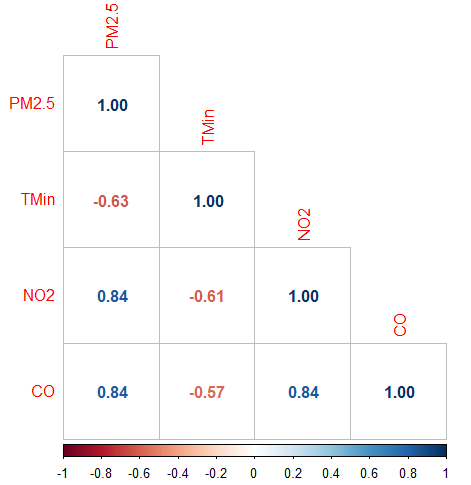
**correlacion <- cor(base\_filtrada)**

**print(correlacion)**



## Matriz de correlación

**corrplot(correlacion, type="lower", method = "number")**



## Coeficiente de determinación del modelo

Adjusted R-squared: 0.7741

## Pruebas de Hipótesis para Correlaciones (𝜌)

Hipótesis ro

**H0: ro = 0**

**H1: ro distinto de 0**

**1. Correlación entre PM2.5 y TMin**

*> cor.test(base$PM2.5, base$TMin)*

t = -10,095, df = 158, p-value < 2.2e-16

**Conclusión**: Existe una correlación negativa moderada significativa entre PM2.5 y TMin, con un valor t de -10,095 y un p-value < 2.2e-16.

**2. Correlación entre PM2.5 y NO2**

*> cor.test(base$PM2.5, base$NO2)*

t = 19,528, df = 158, p-value < 2.2e-16

**Conclusión**: Existe una correlación positiva fuerte significativa entre PM2.5 y NO2, con un valor t de 19,528 y un p-value < 2.2e-16.

**3. Correlación entre PM2.5 y CO**

*> cor.test(base$PM2.5, base$CO)*

t = 19,456, df = 158, p-value < 2.2e-16

**Conclusión**: Existe una correlación positiva fuerte significativa entre PM2.5 y CO, con un valor t de 19,456 y un p-value < 2.2e-16.

**4. Correlación entre TMin y NO2**

*> cor.test(base$TMin, base$NO2)*

t = -9,6113, df = 158, p-value < 2.2e-16

**Conclusión**: Existe una correlación negativa moderada significativa entre TMin y NO2, con un valor t de

-9,6113 y un p-value < 2.2e-16.

**5. Correlación entre TMin y CO**

*> cor.test(base$TMin, base$CO)*

t = -8,7781, df = 158, p-value = 2.572e-15

**Conclusión**: Existe una correlación negativa moderada significativa entre TMin y CO, con un valor t de

-8,7781 y un p-value de 2.572e-15.

**6. Correlación entre NO2 y CO**

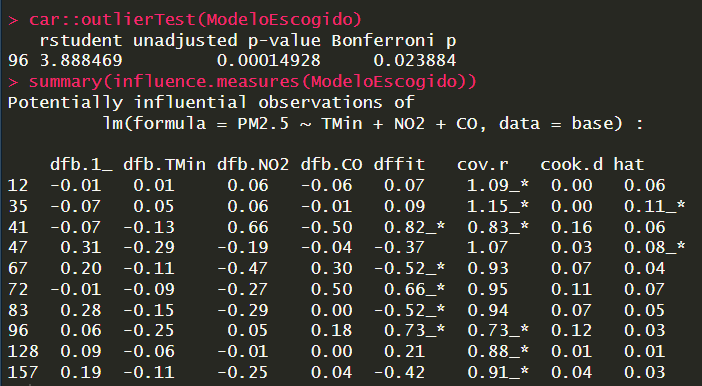
*> cor.test(base$NO2, base$CO)*

t = 19,843, df = 158, p-value < 2.2e-16

**Conclusión**: Existe una correlación positiva fuerte significativa entre NO2 y CO, con un valor t de 19,843 y un p-value < 2.2e-16.

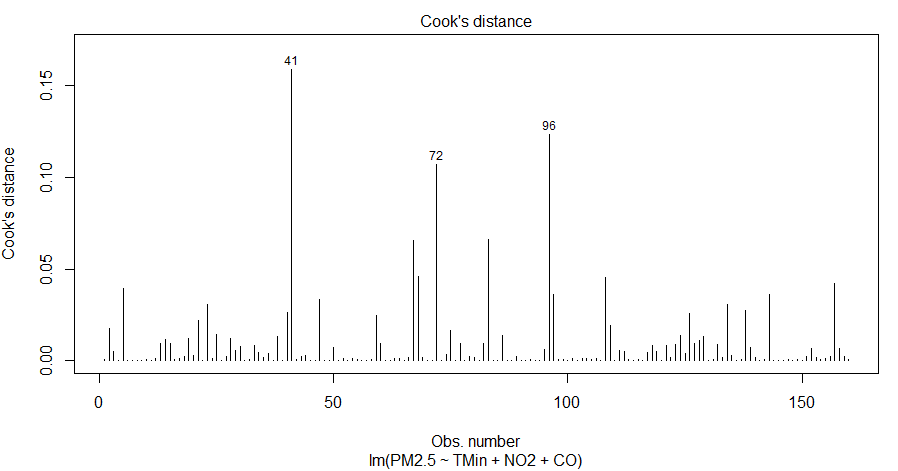
**Comentario general:** en todas las pruebas, los valores p son extremadamente bajos (mucho menores a 0,05), lo que indica que todas las correlaciones encontradas son estadísticamente significativas.

## Valores atípicos o influyentes



## Cook

**plot(ModeloEscogido,4)**



## Leverage

**plot(ModeloEscogido,5)**

