

# Análisis de Clasificación

-Renta Per Cápita Municipios Madrid-



Sergio Casares Fernández  
Beatriz Quevedo Gómez  
Gonzalo Carretero Ursúa

## 1 INTRODUCCIÓN

- OBJETIVO
- VARIABLES

## 2 ANALISIS DISCRIMINANTE

- ANÁLISIS DISCRIMINANTE LINEAL
- PRUEBAS DE NORMALIDAD MULTIVARIADA
- MATRIZ DE CONFUSIÓN Y PORCENTAJE DE ACIERTO

## 3 INTERPRETACIÓN

- MATRIZ DE CORRELACIONES
- GRÁFICOS DE PARTICIONES

# Introducción

## Objetivo

Clasificación de los municipios de Madrid en función del cuartil

## Renta per capita

- 1) Renta baja: 0- 15.000
- 2) Renta medio-baja: 15.000 - 20.000
- 3) Renta medio-alta: 20.000 - 25.000
- 4) Renta alta: 25000 - En adelante

Fuente: Instituto de estadística de la Comunidad de Madrid (2014)

# Variables Explicativas

Autobus	Farmacias	Centros salud	Consultorios	Colegios
Vivienda principal	Segunda vivienda	Viviendas	Edificios	Poblacion
<b>Envejecimiento</b>	Progresividad	Dependencia	Reemplazamiento	<b>Índice de Feminidad</b>
Natalidad	Fecundidad	Mortalidad	Nupcialidad	<b>Crecimiento vegetativo</b>
Paro	Paro juvenil	Contratos temporales	<b>SS</b>	Agricultura y Ganadería
<b>Industria</b>	Construccion	Sector servicios	Servicios financieros	Otros servicios
<b>Km2</b>	Altitud	Distancia km capital	Volumen agua	Turismos
PP	PSOE	IU	UPyD	<b>Otros Partidos</b>

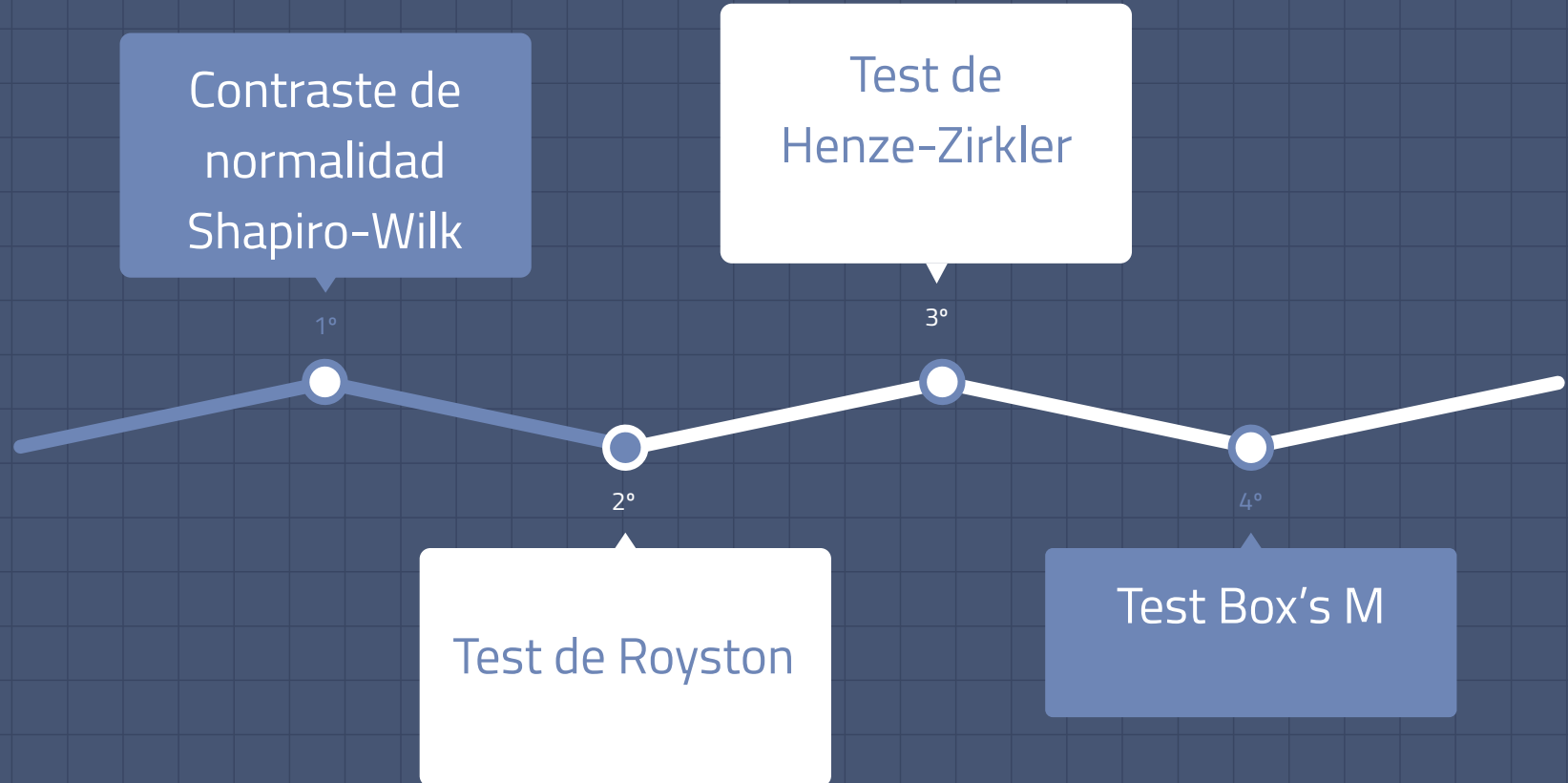
Fuente: Instituto de estadística de la Comunidad de Madrid (2014)

# Análisis Discriminante

El Análisis Discriminante Lineal (LDA) es un método de **clasificación** de variables cualitativas en el que dos o más grupos son conocidos a priori y nuevas observaciones se clasifican en uno de ellos **en función de sus características**.



# Pruebas de normalidad multivariada





# Análisis Discriminante Cuadrático

```
```{r}
modelo <- qda(renta_cat ~
               ss_1m +
               g_envejecimiento +
               t_feminidad +
               mineria_ind_energia +
               km2 +
               otros_partidos +
               crecimiento_vegetativo_1m,
               data = mad)
```
```



# Matriz de Confusión

|             |                   | PREDICCIÓN |                  |                  |            |
|-------------|-------------------|------------|------------------|------------------|------------|
|             |                   | Renta baja | Renta medio-baja | Renta medio-alta | Renta alta |
| OBSERVACIÓN | Renta baja        | 60         | 15               | 3                | 0          |
|             | Renta medio-baja  | 30         | 15               | 0                | 3          |
|             | Renta medio- alta | 0          | 4                | 18               | 4          |
|             | Renta alta        | 2          | 1                | 2                | 25         |

79,8%

# Interpretación

A large, light blue number 3 is positioned on the right side of the slide. The background is a dark blue grid with a silhouette of a bar chart at the bottom. The word 'Interpretación' is written in white in the upper center.

3

# Interpretación - Matriz de correlación

**Variable Dependiente**

**Variables Independientes**

**renta\_cat**



**pib\_pc**

**Min. : 9215**  
**1st Qu. : 13644**  
**Median : 16585**  
**Mean : 19940**  
**3rd Qu. : 21243**  
**Max. : 96832**

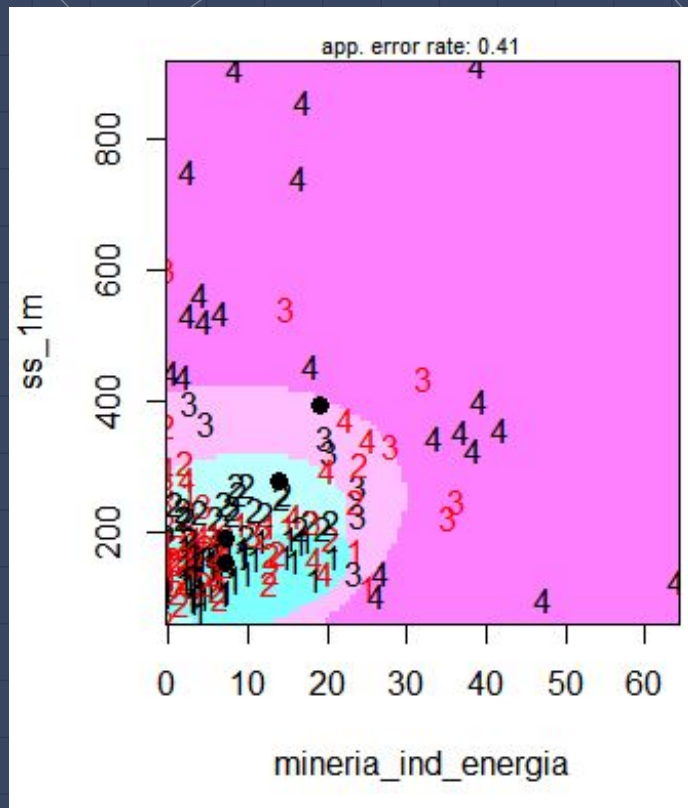
**ss\_1m**

**0.6140**

**mineria\_ind\_energia**

**0.5040**

# Interpretación - Gráfico de Partición



# Conclusión

## Elección del Modelo

Shapiro - Wilk

Test de Royston

Henze - Zirkler

Box's M

QDA

## Variables Utilizadas

### Árbol de decisión

renta\_cat

ss\_1m

km2

g\_envejecimiento

otros\_partidos

t\_feminidad

crec\_vegetativo\_1m

min\_ind\_energia

% de acierto

79,8 %

# GRACIAS!

