



Islas Canarias  
Del 15 al 19 de noviembre de 2021



Consejería de Transformación  
Económica, Industria,  
Conocimiento y Universidades

Instituto de Estadística y  
Cartografía de Andalucía

# **XXI Jornadas de Estadística de las Comunidades Autónomas**

## **Un caso de uso de R aplicado al web scraping en estadísticas de inserción laboral**

Víctor Montañés Cobo

# Estadísticas de egresados



Formación Profesional

Centros Públicos y privados  
de Andalucía



Sistema Séneca

Consejería de Educación



Universidades

Centros Públicos de  
Andalucía (incluyendo  
campus de Universidad de  
Granada)

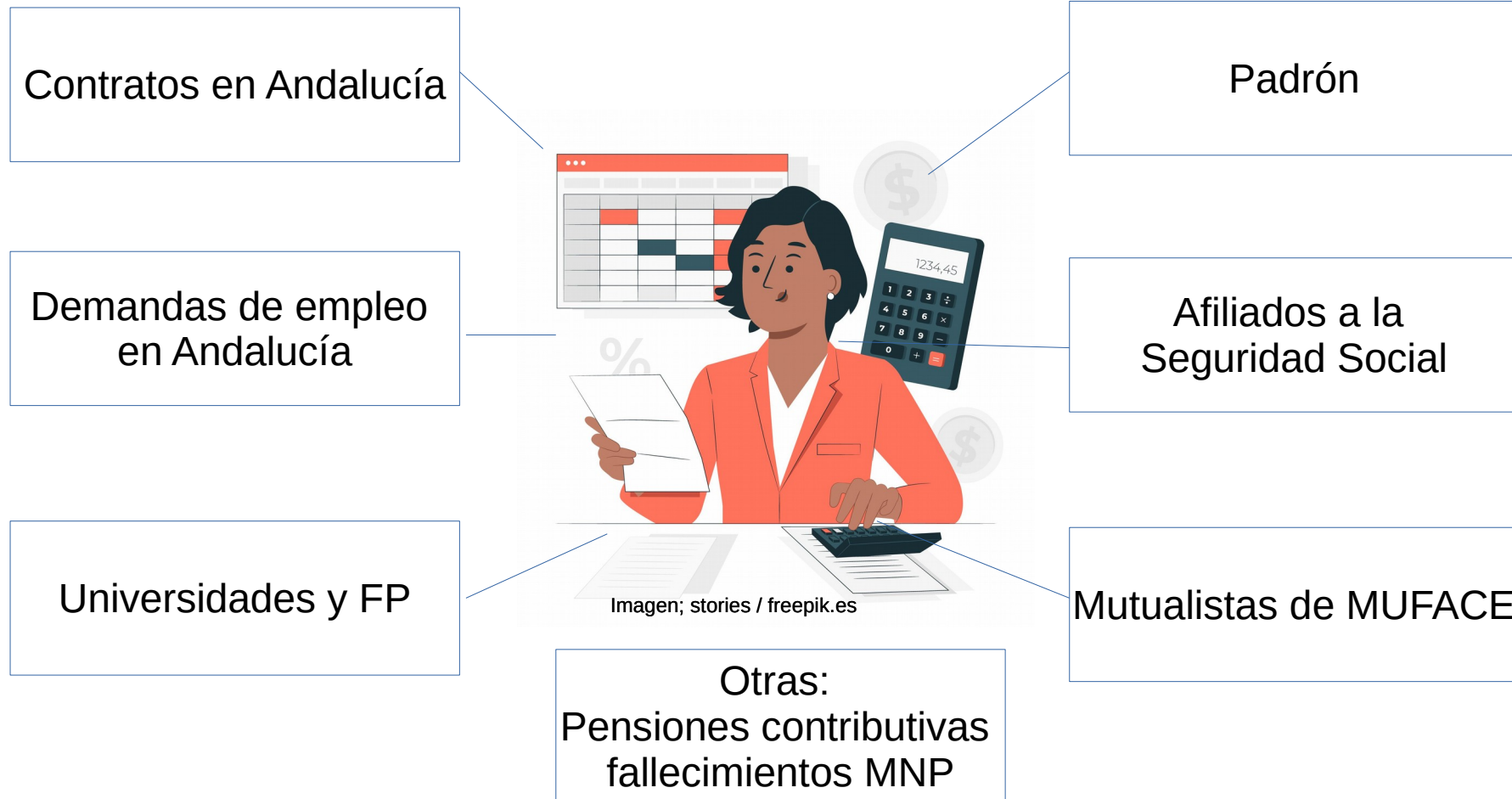


Cada universidad

Información compatible con  
la remitida a SIIU

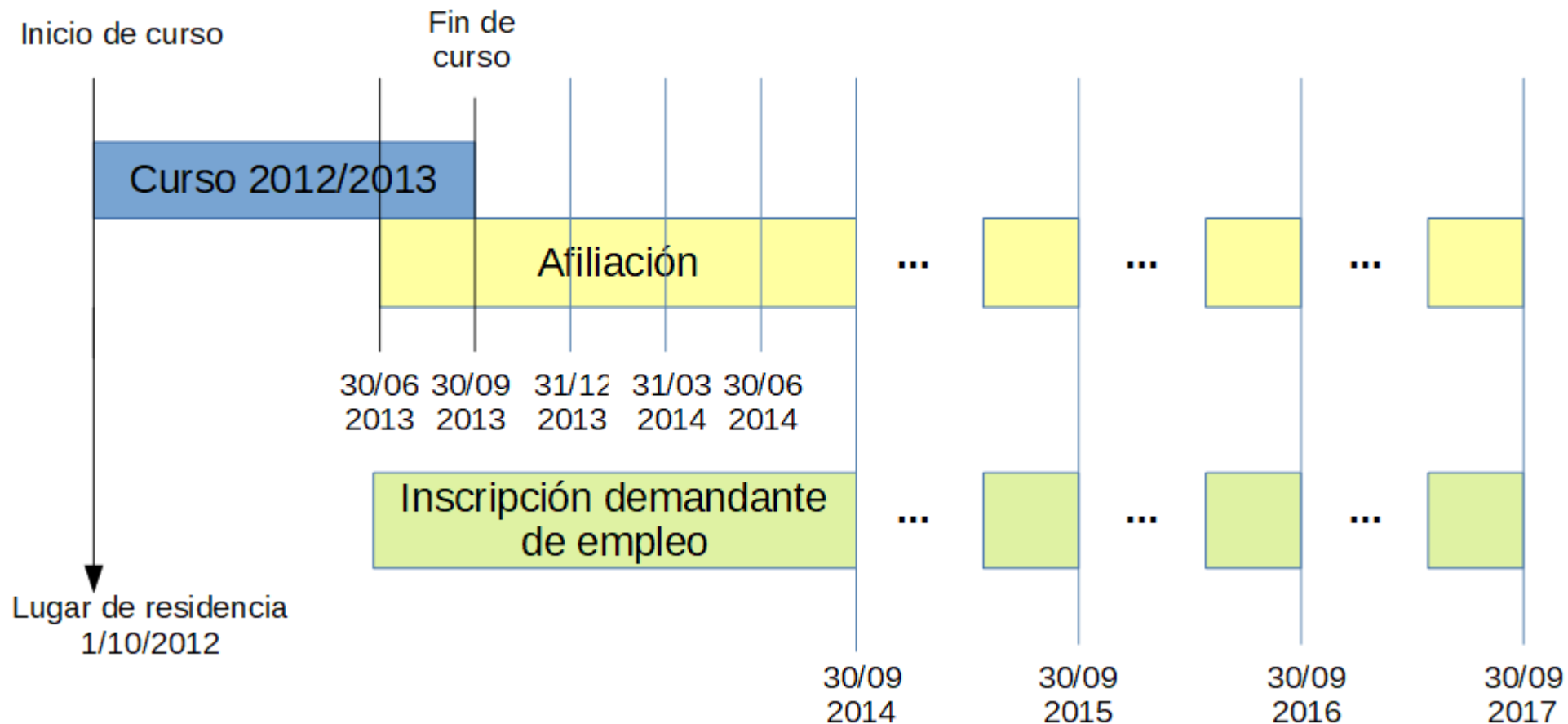
Imágenes: iconicbestiary / freepik.es

# Fuentes de información

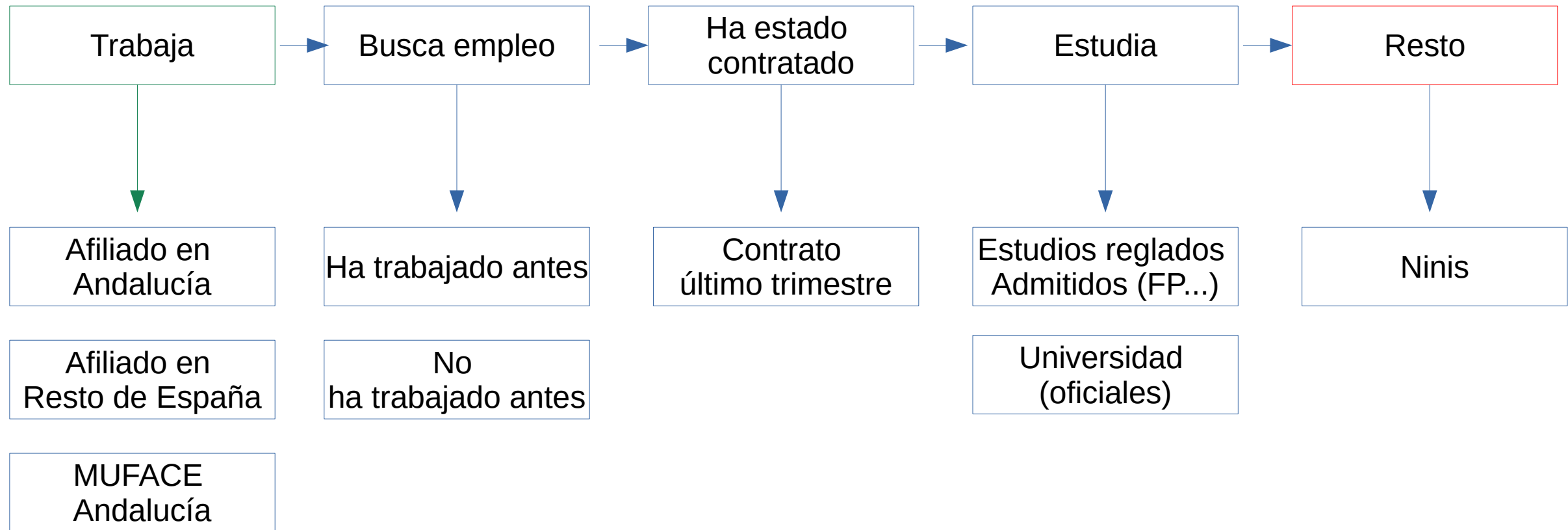


# Seguimiento

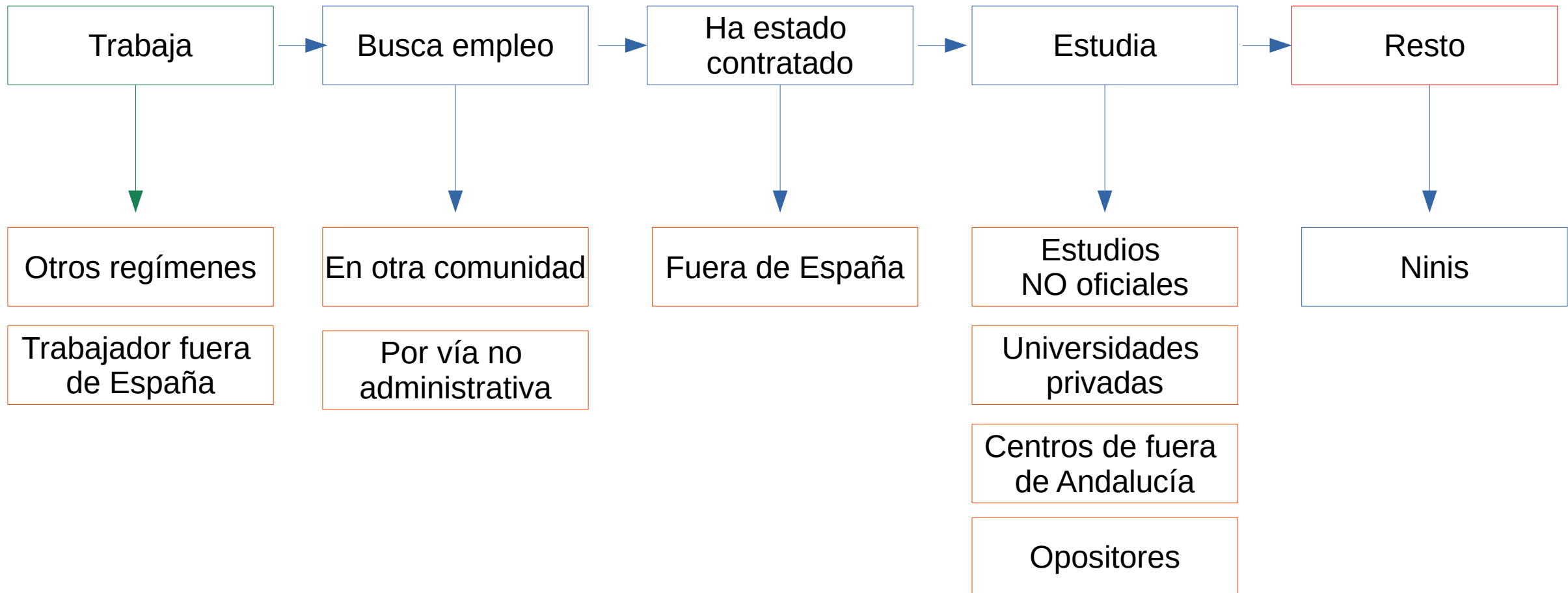
Se asume que el egreso se produce 1 día antes del inicio del siguiente curso.



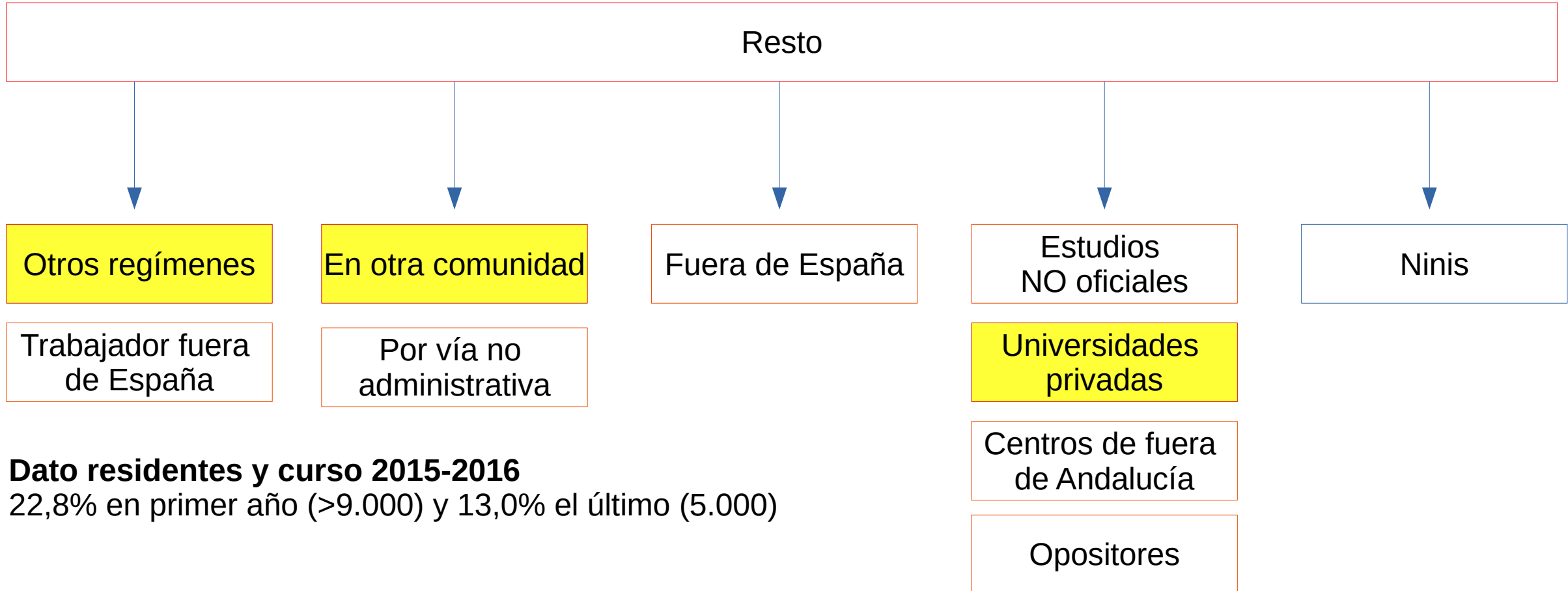
# Clasificación de los individuos



# Lagunas en la clasificación de los individuos

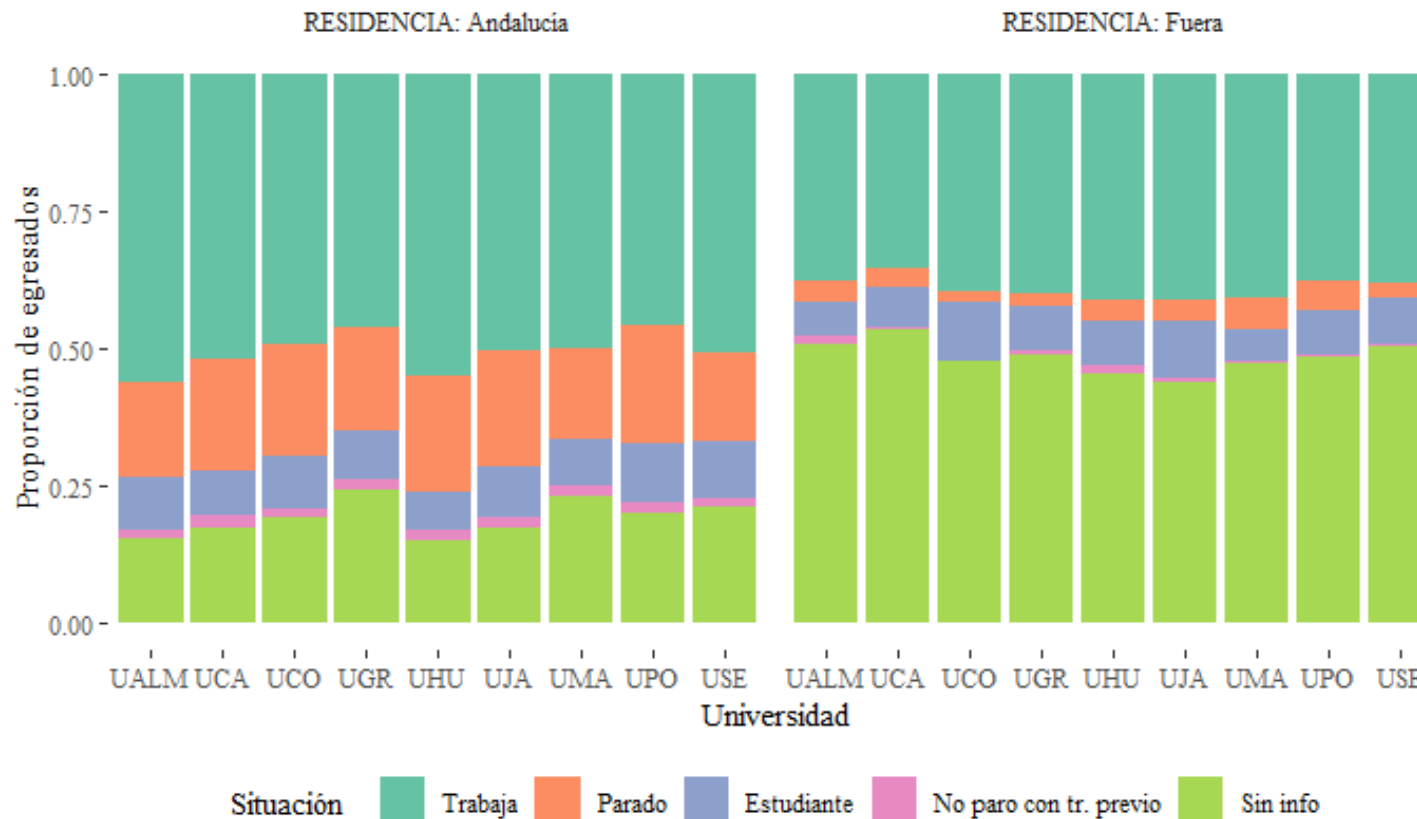


# ¿Quiénes son el resto?



# Lagunas de información / “Resto”

Distribución de egresados por situación, universidad y residencia



Curso 2018-2019. Trimestre 3, Año +1

## ¿Qué es la empleabilidad?

- Capacidad de absorción del mercado laboral andaluz
- Proporción con empleo conocido
- Proporción con empleo de los que hay muestras de querer incorporarse al mercado laboral
- Proporción con empleo compatible en categoría con su titulación
- Capacidad de absorción de los estudiantes andaluces





Islas Canarias  
Del 15 al 19 de noviembre de 2021

# Recuperación de información con R



Consejería de Transformación  
Económica, Industria,  
Conocimiento y Universidades

Instituto de Estadística y  
Cartografía de Andalucía

# APIS, recuperación manual y web scraping

Homer Simpson ahora se llama  
Max Power.



“Homer al máximo.  
Los Simpson.  
Episodio 216.  
Fox”

# Información previa disponible

## Para residentes o ex-residentes

- Documento de identidad
- Fecha de nacimiento
- Nombre y apellidos
- Universidad y titulación de la que egresa

## Para personas que nunca han residido en Andalucía

- Documento de identidad
- Universidad y titulación de la que egresa



Imagen: wikipedia

# ¿Dónde buscar?

Fuentes oficiales accesibles:

Boletines oficiales:



- Nombramientos
- Admitidos y notas de procesos selectivos

TESEO



Imagen: katemangostar / freepik.es

Fuentes privadas:

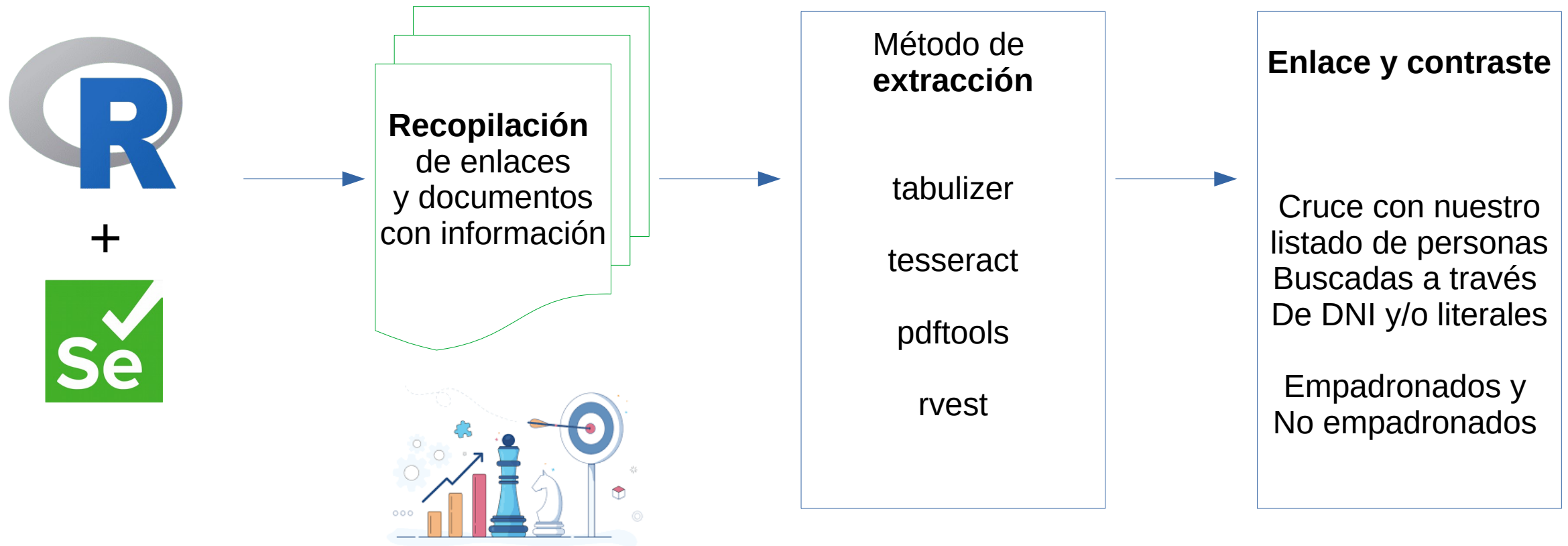
**LinkedIn**

- Búsqueda de empleo no oficial
- Trabajo fuera de España
- Contratos temporales fuera de Andalucía
- Continuación de estudios

**R<sup>G</sup> ResearchGate**

- Trabajo en el extranjero
- Continuación de estudios

# Estrategia 1. Boletines oficiales



# Boletines oficiales



Apartado nombramientos:  
Personal de ISFAS

Trabajadores  
(se desconoce el  
centro de destino)



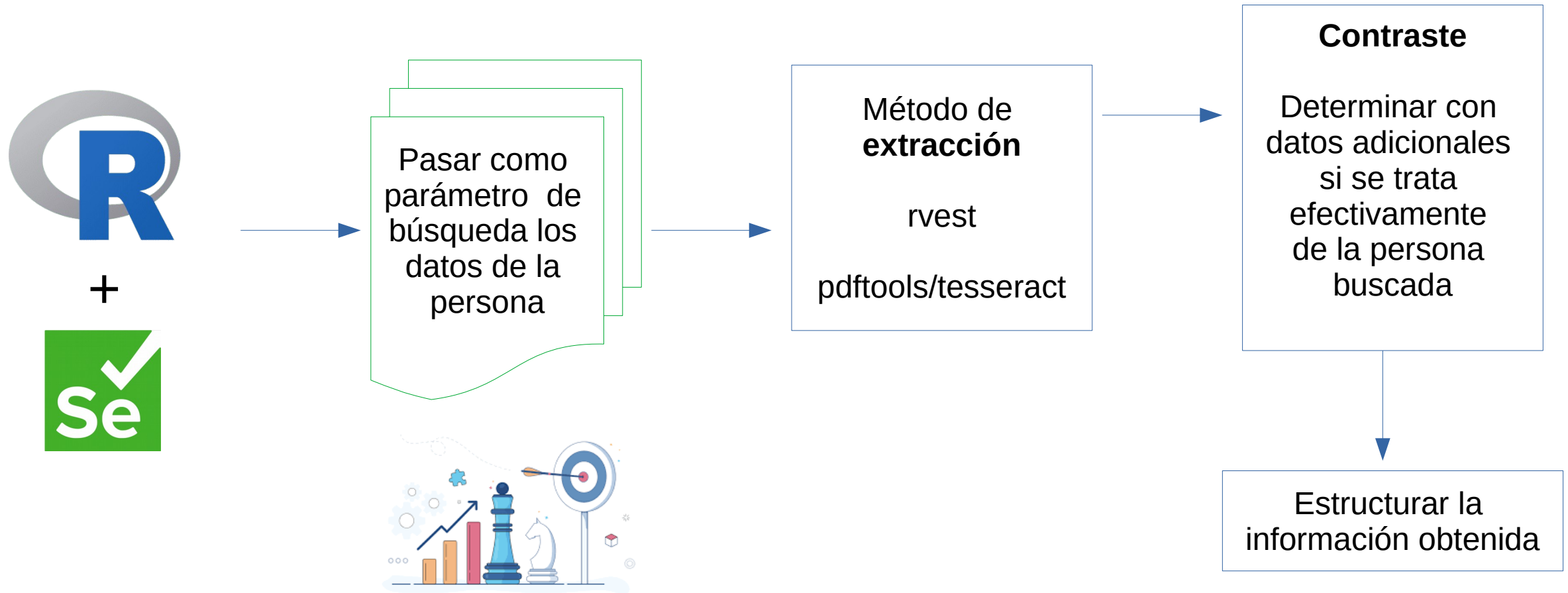
Convocatorias de empleo:  
Oposiciones  
Educación Primaria

Estudiantes / opositores

## Hallazgos:

- Pocas recuperaciones empleados de ISFAS. Algo más de 100
- **Identificación de un patrón estacional** de continuación de estudios en Graduados en Educación Primaria
- Un 21,1% de los graduados en E.P. calificados como resto, se presentan al examen. Un 10,1% además de presentarse, aprueban algún examen.

## Estrategia 2. Teseo, LinkedIn, ResearchGate







## Experiencia



### Instituto de Estadística y Cartografía de Andalucía

18 años y 1 mes

Asesor técnico de proyectos estadísticos  
sept 2017 – actualidad · 4 años y 1 mes  
Avenida Leonardo Da Vinci nº21, Sevilla

## Educación



### Universidad de Granada

Máster en Estadística Aplicada, Estadística  
2011 – 2012

Trabajo de fin de Máster: Distribuciones tipo fase.



### Víctor Montañés Cobo

Asesor técnico de proyectos estadísticos en Instituto de Estadística y Cartografía de Andalucía

Sevilla y alrededores · [Información de contacto](#)

94 contactos

## Hallazgos

- Alrededor de  $\frac{1}{3}$  tienen perfil
- El 40% de los muestreados saldrían del grupo resto
- De los estudiantes,  $\frac{1}{3}$  estudiaban enseñanzas no regladas: Escuelas de Negocio





## Current affiliation

**Universidad de Sevilla** ←

Location  
Sevilla, Spain ←

Department  
Department of Building Structures and  
Geotechnical Engineering

Position  
Lecturer

Time period  
Feb 2019 - Present ←



*buildings*



Article

## Action protocols for seismic evaluation of structures and damage restoration of residential buildings in Andalusia (Spain): "IT-Sismo" APP

Emilio J. Mascort-Albet <sup>1,\*</sup>, Jacinto Canivell <sup>2</sup>, Antonio Jaramillo-Morilla <sup>1</sup>, Rocío Romero-Hernández <sup>1</sup>, Jonathan Ruiz-Jaramillo <sup>3</sup> and Cristina Soriano-Cuesta <sup>1</sup>

<sup>1</sup> Department of Building Structures and Soil Engineering. Higher Technical School of Architecture. Universidad de Sevilla, 41012 Sevilla, Spain; jarami@us.es (A.J.-M.); rocirome@us.es (R.R.-H.); csoriano@us.es (C.S.-C.)

<sup>2</sup> Department of Architectural Construction II. Higher Technical School of Building Engineering. Universidad de Sevilla, 41012 Sevilla, Spain; jacanivell@us.es

<sup>3</sup> Department of Art and Architecture, Architectural Construction's Chair. Higher Technical School of Architecture. Universidad de Málaga, 29071 Málaga, Spain; jonaruizjara@uma.es.

\* Correspondence: emascort@us.es; Tel.: +34- 954-556-662

Received: 22 March 2019; Accepted: 25 April 2019; Published: 29 April 2019

Mucho arroz para tan poco pollo

# TESEO

## Tesis doctorales: TESEO

### Consulta de la Base de datos de Tesis Doctorales (TESEO)

Rellene los campos del siguiente formulario con las palabras o frases que conozca de los documentos

Búsqueda Simple
  Entrar como usuario registrado
  Registrarme como Doctoran

**Buscar Tesis**

con todas las palabras
  con la frase exacta
  con alguna de las palabras
  sin ninguna de las palabras

...en el

**Otros Campos de Búsqueda**

Autor:   
 NIF/NIE/Pasaporte:   
 Director:   
 Universidad: -- Universidades --  
 Departamento: -- 0 Departamentos --  
 Curso académico: Desde:  /  Hasta:  /

#nif  Clear (1)

## Tesis doctorales: TESEO

### Resultado de la búsqueda

Mostrando ficha 1 de 1 fichas de tesis seleccionadas.

- > **Título:** ACTUALIZACIÓN DE MARCOS INPUT-OUTPUT A TRAVÉS DE MÉTODOS DE PROYECCIÓN: E EN TABLAS DE ORIGEN Y DE DESTINO A PRECIOS BÁSICOS DE VARIOS PAÍSES DE LA UNIÓN EURO
- > **Autor:** VALDERAS JARAMILLO, JUAN MANUEL
- > **Universidad:** Universidad de Sevilla
- > **Departamento:** Economía aplicada I
- > **Fecha de Lectura:** 22/01/2016
- > **Programa de doctorado:** Economía Aplicada
- > **Dirección:**
  - > RUEDA CANTUCHE, JOSE MANUEL (Director)
  - > OLMEDO FERNANDEZ, ELENA (Director)
- > **Tribunal:**
  - > ROBLES TEIJEIRO, LUÍS (presidente)
  - > Fernández Vázquez, Esteban (secretario)
  - > ESCOT MANGAS, LORENZO (vocal)
  - > DUARTE PAC, Mª ROSA (vocal)
  - > GRAU CARLES, MARIA DEL PILAR (vocal)
- > **Descriptor:**
  - > INPUT-OUTPUT
  - > ESTADISTICA ECONOMICA
  - > PROYECCION ECONOMICA
  - > SISTEMAS ECONOMICOS
- > **Localización:** UNIVERSIDAD DE SEVILLA, SERVICIO DE DOCTORADO
- > **Marcador:** <https://www.educacion.gob.es/teseo/mostrarRef.do?ref=1184631> Añadir a favoritos.
- > **Resumen:**

En este trabajo se estudian en profundidad los métodos de proyección EURO y SUT-RAS para



Islas Canarias  
Del 15 al 19 de noviembre de 2021

# Conclusiones



Consejería de Transformación  
Económica, Industria,  
Conocimiento y Universidades

Instituto de Estadística y  
Cartografía de Andalucía



@vmontanes

# Conclusiones

- **Carencias irresolubles** en fuentes administrativas
- Veracidad del Scraping de **documentos oficiales**
- Limitación, sesgos y sostenibilidad del uso de las fuentes privadas
- LinkedIn se muestra como la fuente más útil
- ...y si quieres tener visitas en tu perfil de LinkedIn, haz scraping.



Islas Canarias  
Del 15 al 19 de noviembre de 2021

# ¡Gracias!



Consejería de Transformación  
Económica, Industria,  
Conocimiento y Universidades

Instituto de Estadística y  
Cartografía de Andalucía



@vmontanes