

# Hola!



**WOMEN IN DATA SCIENCE**  
@ STANFORD UNIVERSITY

---

## Soy Vilma Romero

Docente en Universidad Nacional de Ingeniería

vromeror@uni.edu.pe

<https://vilmaromero.github.io>

@vilsurr

5 Marzo, 2018



# Tabla de Contenidos

1. Acerca de mí
2. Proyectos principales realizados
3. Capacitadora, Instructora y Docente
4. Herramientas y Recursos
5. Proyectos en marcha
6. Proyectos a futuro

## Acerca de mí

---

# Acerca de mí

**Desde 2016** **Docente de Estadística**

Universidad Nacional de Ingeniería



**2014 - 2015** **MSc. en Estadística**

Lancaster University, UK

**2014** **Ingeniera Estadística**

Universidad Nacional de Ingeniería

**2008 - 2012** **BSc. en Ingeniería Estadística**

Universidad Nacional de Ingeniería

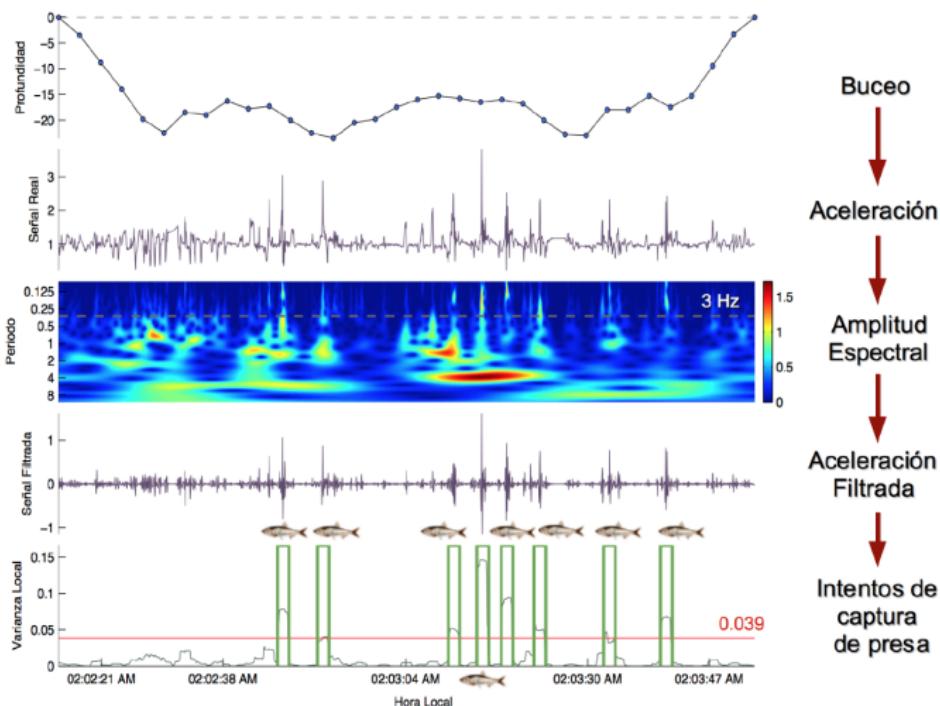
## **Proyectos principales realizados**

---

## Identificación de intentos de captura de presa usando datos de acelerometría en los Lobos Marinos Finos Sudamericanos

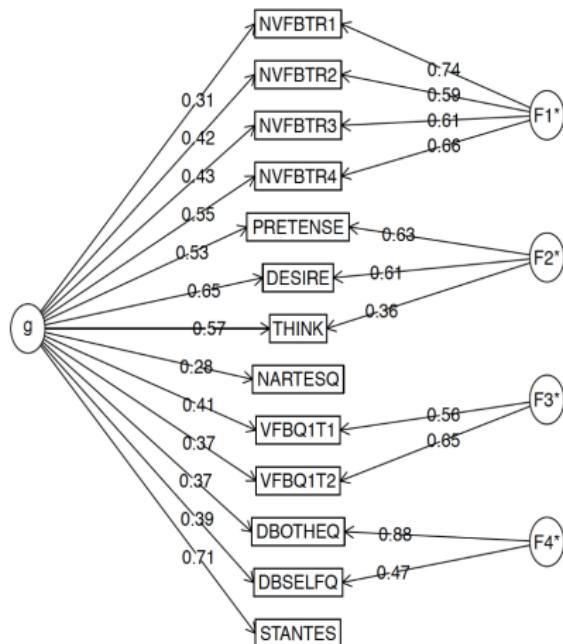


Se equipó a 5 lobas marinas durante etapa de reproducción en la colonia Punta San Juan (Ica) en el 2011.



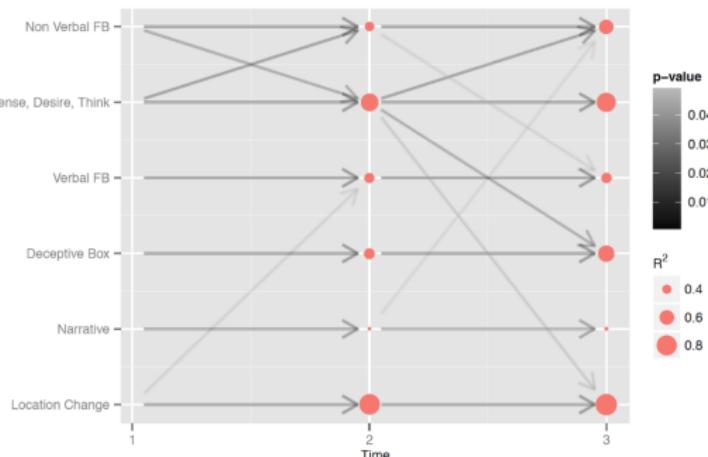


## Longitudinal Multidimensional Item Response Modelling in Preschool Children's Mental State Understanding



### Participants

86 British children (Female = 41, Male = 45) from different preschools and day nurseries located in Northern Lancashire.  
Age: Between 30 and 33 months when recruited.



Path Diagram of Causality - Model AR(1).

The p-values have not been adjusted for multiple comparison.

**Capacitadora, Instructora y Docente**

---

## 1. Minería de Datos con R

- Reducción de Dimensionalidad
- Métodos Supervisados y No Supervisados
- Market Basket Analysis
- Análisis de Datos Longitudinales

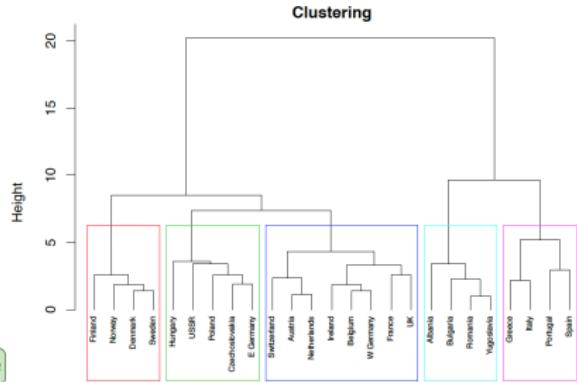
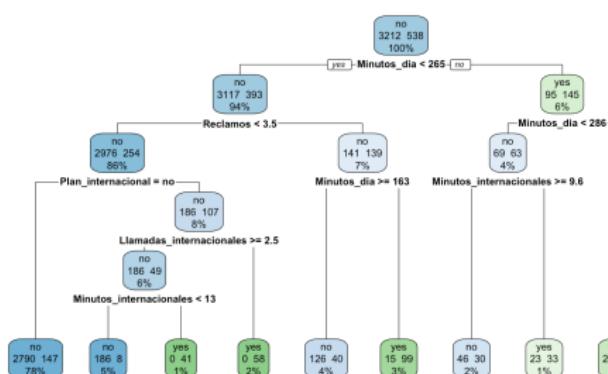


## 2. Análisis de Datos con R

## 3. Reportes Dinámicos con R

## 4. Introducción a Latex

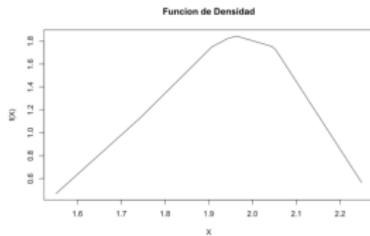
LATEX



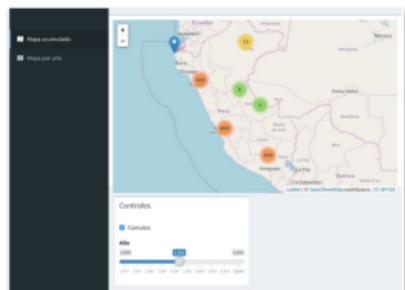
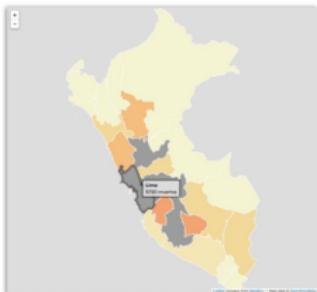
## Feria de Proyectos 2017-I



Función de Densidad y Acumulativa de una Distribución de Weibull

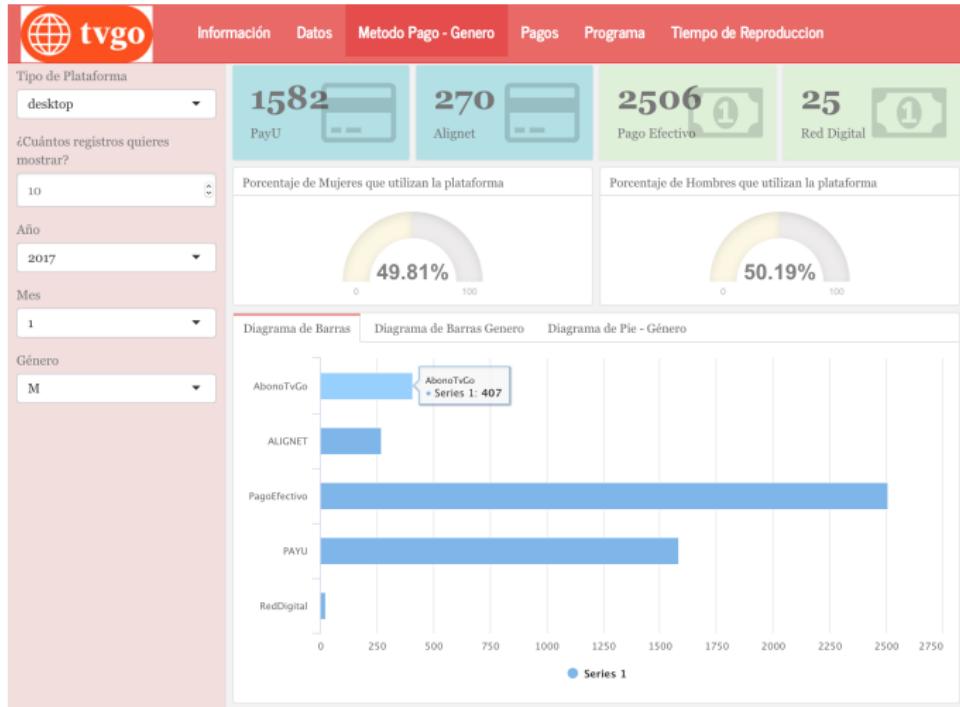


## Feria de Proyectos 2017-II



# Diplomado PEBIBA - SEUPROS - UNI

- Herramientas avanzadas para visualización de datos en R  
(Flexdashboard, Shiny, ggplot2)



## **Herramientas y Recursos**

---

# Data Scientist - Cursos en Línea - Repositorio de Datos

## MODERN DATA SCIENTIST

Data Scientist, the sexiest job of the 21th century, requires a mixture of multidisciplinary skills ranging from an intersection of mathematics, statistics, computer science, communication and business. Finding a data scientist is hard. Finding people who understand who a data scientist is, is equally hard. So here is a little cheat sheet on who the modern data scientist really is.

### MATH & STATISTICS

- ★ Machine learning
- ★ Statistical modeling
- ★ Experiment design
- ★ Bayesian inference
- ★ Supervised learning: decision trees, random forests, logistic regression
- ★ Unsupervised learning: clustering, dimensionality reduction
- ★ Optimization: gradient descent and variants



### PROGRAMMING & DATABASE

- ★ Computer science fundamentals
- ★ Scripting language e.g. Python
- ★ Statistical computing packages, e.g. R
- ★ Databases: SQL and NoSQL
- ★ Relational algebra
- ★ Parallel databases and parallel query processing
- ★ MapReduce concepts
- ★ Hadoop and Hive/Pig
- ★ Custom reducers
- ★ Experience with xaaS like AWS

### COMMUNICATION & VISUALIZATION

- ★ Able to engage with senior management
- ★ Story telling skills
- ★ Translate data-driven insights into decisions and actions
- ★ Visual art design
- ★ R packages like ggplot or lattice
- ★ Knowledge of any of visualization tools e.g. Flare, D3.js, Tableau

### DOMAIN KNOWLEDGE & SOFT SKILLS

- ★ Passionate about the business
- ★ Curious about data
- ★ Influence without authority
- ★ Hacker mindset
- ★ Problem solver
- ★ Strategic, proactive, creative, innovative and collaborative

## Cursos en línea



## Repositorio de datos



## Google Trends Datastore



Imágenes tomadas de Parvaneh Shafiei

# Herramientas y Recursos

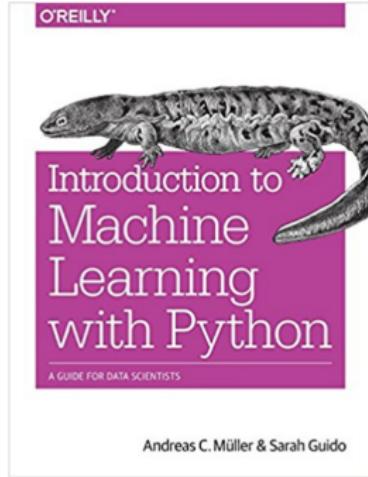
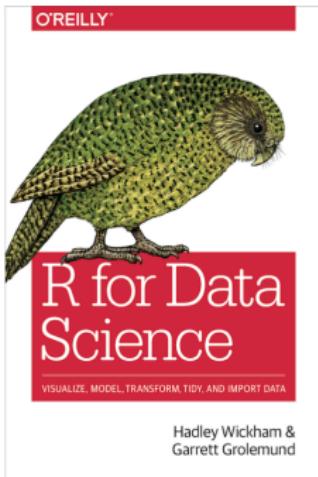
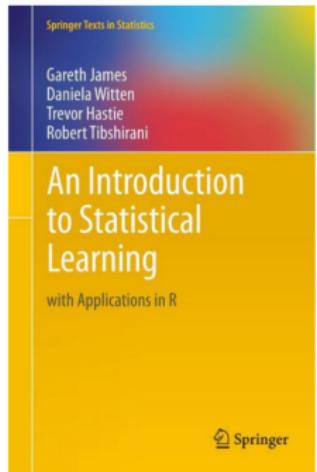


Imagen tomada de Parvaneh Shafiee  
R-Ladies Milan

R-bloggers, Analytics Vidhya, KDnuggets, Simply Statistics, ...



## **Proyectos en marcha**

---



Es una organización global que promueve la **participación de mujeres** en la comunidad **R** a través de **meetups** en un ambiente **amigable** y **colaborativo**.



**Facebook:** /RLadiesLima    **Twitter:** @RLadiesLima

## **Proyectos a futuro**

---

En un futuro próximo:

- Realizar un PhD. en Estadística o Ciencia de Datos.
- Aplicar visualización de datos para el conocimiento estadístico.
- Crear un nexo entre el mundo empresarial y académico mediante consultoría.

# Muchas gracias!!

