

Trabajo Práctico 1

[75.06 / 95.58] Organización de Datos
Primer Cuatrimestre de 2018

Alumno	Padrón	E-mail
Soro, Lucas Gustavo	95665	lugusor@gmail.com
Hazan, Pablo Nehuén	96522	phazan.fiuba@gmail.com
Núñez Leyes, Javier Damián	94455	javier.nunezleyes@gmail.com
Montes, Marcelo	81397	mdmontes@gmail.com

Índice

1_ Introducción

2_ Pre-procesamiento del set de datos

1_ Introducción

El informe expone los resultados del análisis exploratorio de los datos del registro histórico de avisos de búsquedas laborales en <https://www.zonajobs.com.ar>

Link al set de datos:

https://drive.google.com/file/d/1K4uRag5nmGtfuvzyJV9RL_73lzsh_iTO/view?usp=sharing

Link al repositorio con los notebook en GitHub:

<https://github.com/javierleyes/OrganizacionDeDatos>

2_ Pre-procesamiento del set de datos

Para cada set de datos analizamos los tipos de datos, la presencia de valores nulos o no validos, los identificadores usados y la posibilidad de relacionar los datos entre distintos sets con estos identificadores.

Se encontraron casos de fechas nulas y con formato no valido (personas menores de edad y mayores a 100 años) por ser una cantidad despreciable se descartaron los casos irregulares.

Para el caso de las columnas donde existian muy pocos datos validos para usar, se opto por descartar la columna.