

NBA DATA WAREHOUSE

BASES DE DATOS AVANZADAS

- Mario Pérez Sánchez-Montañez
- David Camuñas Sánchez



ÍNDICE

- Fuente de datos
- Obtención de estadísticas
- Realización de consultas
- Modelo de regresión lineal



FUENTE DE DATOS

- Los datos utilizados para la creación del almacén de datos se han obtenido de www.basketball-reference.com



Fuente: <https://www.basketball-reference.com>

CREACIÓN DEL ALMACÉN DE DATOS

- Obtención de los datos aplicando técnicas de *web scraping* con **Python**
- Creación de *dataframes* con **pandas**
- Creación de la base de datos



Fuente: <https://www.python.org>

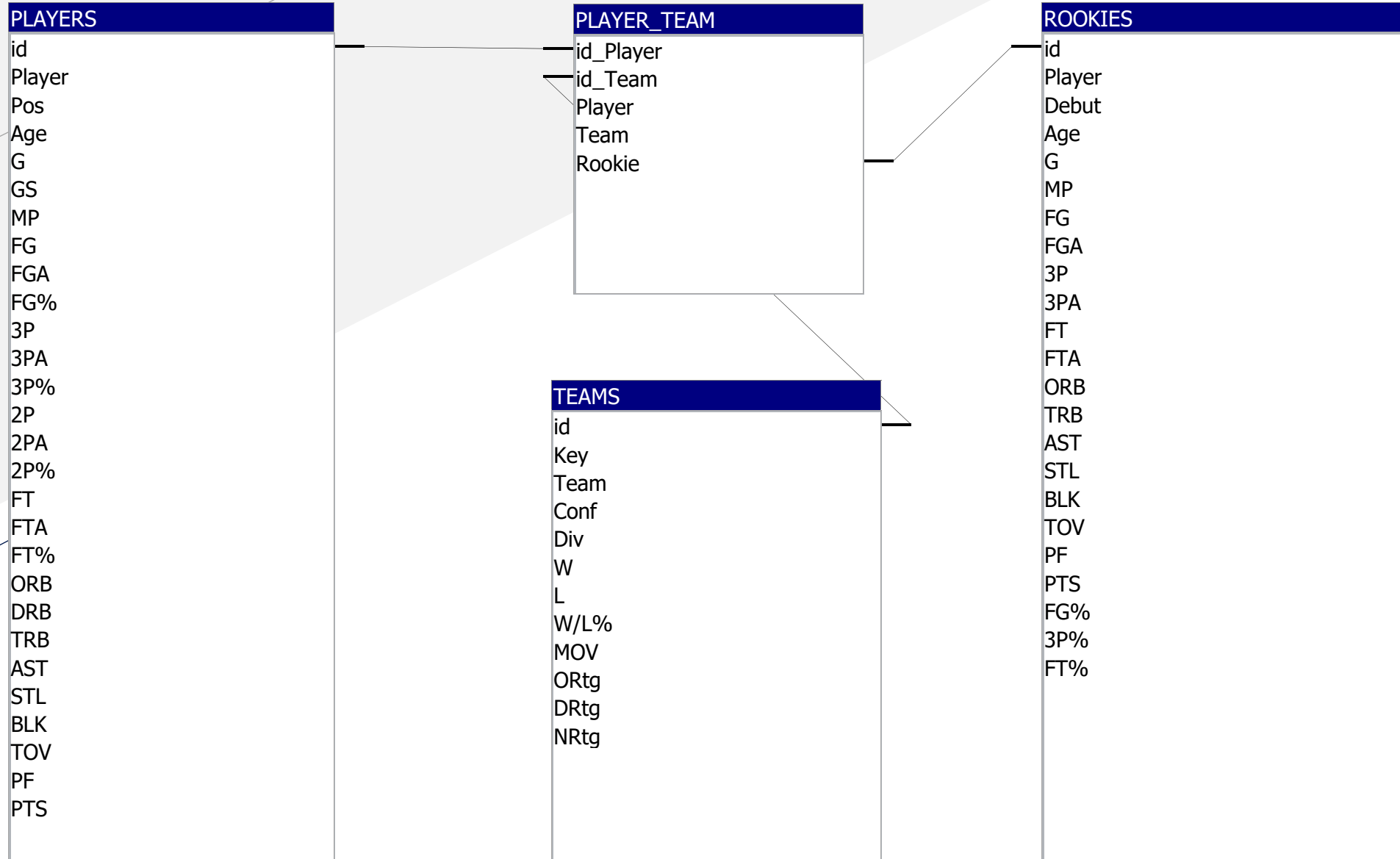


Fuente: <https://pandas.pydata.org>



Fuente: <https://es.wikipedia.org/wiki/SQLite#>

ESQUEMA BASE DE DATOS



OBTENCIÓN DE ESTADÍSTICAS

Realización de consultas (SQLite & Python)

- Mediante la utilización de consultas.
 - Librería SQLite.
- Creación de Dataset.
 - Librería Pandas.
- Creación de gráficos.
 - Librería Matplotlib.

matplotlib

Fuente: <https://matplotlib.org>



Fuente: <https://es.wikipedia.org/wiki/SQLite#>

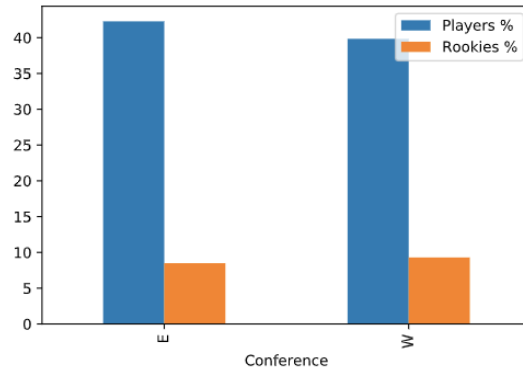
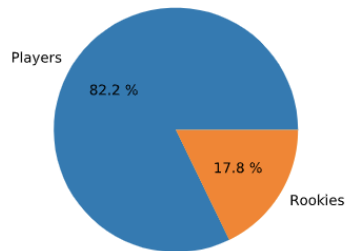


Fuente: <https://pandas.pydata.org>

ALGUNOS EJEMPLOS

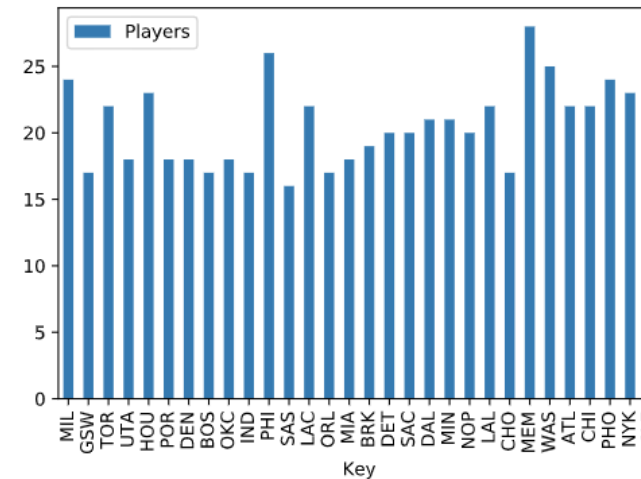
Realización de consultas (SQLite & Python)

- Porcentaje de jugadores *Rookies*.
 - Toda temporada (izquierda) y por conferencias (derecha).



Fuente: *Study.ipynb*

- Nº jugadores de cada equipo.



ALGUNOS EJEMPLOS

Realización de consultas (SQLite & Python)

- Clasificación final de la temporada.
 - Mediante índice W/L (sin *Playoffs*).

	Key	Team	Conf	W	L
0	MIL	Milwaukee Bucks	E	60.0	22.0
1	TOR	Toronto Raptors	E	58.0	24.0
2	GSW	Golden State Warriors	W	57.0	25.0
3	DEN	Denver Nuggets	W	54.0	28.0
4	HOU	Houston Rockets	W	53.0	29.0
5	POR	Portland Trail Blazers	W	53.0	29.0
6	PHI	Philadelphia 76ers	E	51.0	31.0
7	UTA	Utah Jazz	W	50.0	32.0
8	BOS	Boston Celtics	E	49.0	33.0
9	OKC	Oklahoma City Thunder	W	49.0	33.0
10	IND	Indiana Pacers	E	48.0	34.0
11	SAS	San Antonio Spurs	W	48.0	34.0
12	LAC	Los Angeles Clippers	W	48.0	34.0
13	ORL	Orlando Magic	E	42.0	40.0
14	BRK	Brooklyn Nets	E	42.0	40.0

15	DET	Detroit Pistons	E	41.0	41.0
16	MIA	Miami Heat	E	39.0	43.0
17	SAC	Sacramento Kings	W	39.0	43.0
18	CHO	Charlotte Hornets	E	39.0	43.0
19	LAL	Los Angeles Lakers	W	37.0	45.0
20	MIN	Minnesota Timberwolves	W	36.0	46.0
21	DAL	Dallas Mavericks	W	33.0	49.0
22	NOP	New Orleans Pelicans	W	33.0	49.0
23	MEM	Memphis Grizzlies	W	33.0	49.0
24	WAS	Washington Wizards	E	32.0	50.0
25	ATL	Atlanta Hawks	E	29.0	53.0
26	CHI	Chicago Bulls	E	22.0	60.0
27	PHO	Phoenix Suns	W	19.0	63.0
28	CLE	Cleveland Cavaliers	E	19.0	63.0
29	NYK	New York Knicks	E	17.0	65.0

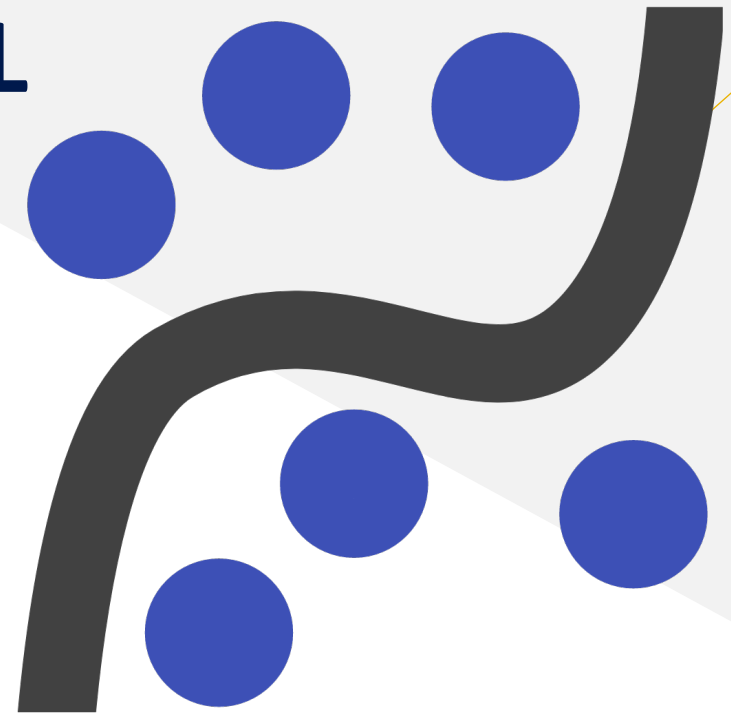
- Mejor equipo de cada conferencia
 - Cantidad de victorias, derrotas, margen de victorias...

	Key	Team	Conf	W	L	W/L%	MOV	ORtg	DRtg	NRtg
0	MIL	Milwaukee Bucks	E	60.0	22.0	0.732	8.87	114.23	105.76	8.47
1	GSW	Golden State Warriors	W	57.0	25.0	0.695	6.46	116.63	110.24	6.39

MODELO DE REGRESIÓN LINEAL

Relación tiros a canasta y puntos

- El objetivo del modelo de regresión lineal es estudiar la relación que existe entre los intentos de tiro y los puntos anotados por jugador
 - $PTS = a + b * TA$
 - $PTS = -5,209429 + 1,000137 * TA$

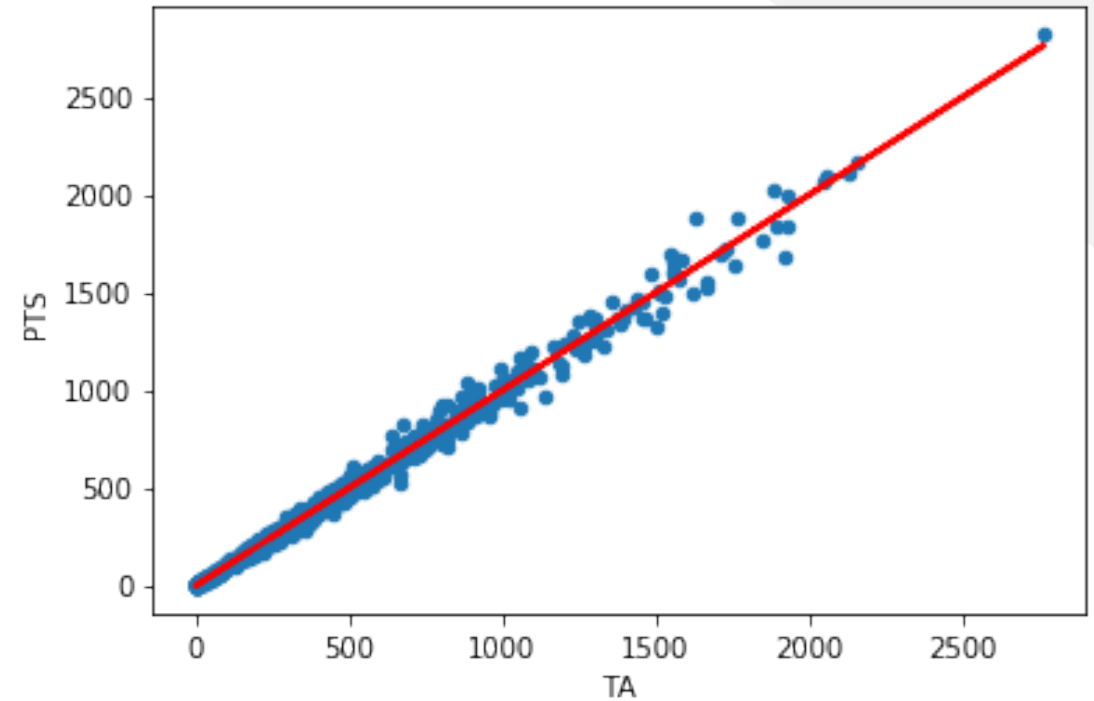


statsmodels

PREDICCIÓN

FR

JUGADOR	PUNTOS REALES	PREDICCIÓN
Álex Abrines	165	164.813847
Quincy Acy	17	22.794404
Jaylen Adams	108	113.806864
Steven Adams	1108	1095.941316



PREGUNTAS

